Analysis, Modeling and Control of the Airport Departure Process

by

Ioannis Simaiakis

Eng. Dipl., National Technical University of Athens (2006)
S.M., Massachusetts Institute of Technology (2009)

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

Doctor of Philosophy in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2013

Author .....
Department of Aeronautics and Astronautics
September 14, 2012

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

Certified by .....
Eric M. Feron
Professor of Aerospace Engineering, Georgia Institute of Technology
Committee Member

Certified by .....
R. John Hansman
Professor of Aeronautics and Astronautics
Committee Member

Certified by .....
Amedeo R. Odoni
Professor of Aeronautics and Astronautics
Committee Member

Accepted by .....
Eytan Modiano
Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Analysis, Modeling and Control of the Airport Departure Process

by

Ioannis Simaiakis

Submitted to the Department of Aeronautics and Astronautics
on September 14, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Aeronautics and Astronautics

Abstract

Increased air traffic demand over the past two decades has resulted in significant increases in surface congestion at major airports in the United States. The overall objective of this thesis is to mitigate the adverse effects of airport surface congestion, including increased taxi-out times, fuel burn, and emissions. The thesis tackles this objective in three steps: The first part deals with the analysis of departure operations and the characterization of airport capacity; the second part develops a new model of the departure process; and the third part of the thesis proposes and tests, both on the field and in simulations, algorithms for the control of the departure process.

The characterization and estimation of airport capacity is essential for the successful management of congestion. This thesis proposes a new parametric method for estimating the departure capacity of a runway system, the most constrained element of most airports. The insights gained from the proposed technique are demonstrated through a case study of Boston Logan International Airport (BOS). Subsequently, the methodology is generalized to the study of interactions among the three main airports of the New York Metroplex, namely, John F. Kennedy International Airport (JFK), Newark Liberty International Airport (EWR) and LaGuardia Airport (LGA). The individual capacities of the three airports are estimated, dependencies between their operations are identified, and the capacity of the Metroplex as a whole is characterized. The thesis also identifies opportunities for improving the operational capacity of the Metroplex without significant redesign of the airspace. The proposed methodology is finally used to assess the relationship between route availability during convective weather and the capacity of LGA.

The second part of the thesis develops a novel analytical model of the departure process. The modeling procedure includes the estimation of unimpeded taxi-out time distributions, and the development of a stochastic and dynamic queuing model of the departure runway(s), based on the transient analysis of $D(t)/E_k(t)/1$ queuing systems. The parameters of the runway service process are estimated using operational data. Using the aircraft pushback schedule as input, the model predicts the expected runway schedule and the takeoff times. It also estimates the expected queuing delay and its variance for each flight, along with the congestion level of the airport, sizes of the departure queues, and the departure throughput. The model is trained using data from EWR in 2011, and is subsequently used to predict taxi-out times at EWR in 2007 and 2010.

The final part of this thesis proposes dynamic programming algorithms for controlling the departure process, given the current operating environment. These algorithms, called *Pushback Rate Control* protocols, predict the departure throughput of the airport, and recommend a rate at which to release pushbacks from the gate in order to control congestion. The thesis describes the design and field-testing of a variant of Pushback Rate Control at BOS in 2011, and the development of a decision-support tool for its implementation. The analysis shows that during 8 four-hour test periods, fuel use was reduced by an estimated 9 US tons (2,650 US gallons), and taxi-out times

were reduced by an average of 5.3 min for the 144 flights that were held at the gate. The thesis concludes with simulations of the Pushback Rate Control protocol at Philadelphia International Airport (PHL), one of the most congested airports in the US, and a discussion of the potential benefits and implementation challenges.

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

Acknowledgments

First and foremost, I would like to thank my advisor, Hamsa Balakrishnan, for her vision, supervision and support during the last five years. I would like to thank Hamsa not only for her guidance in the work presented in this thesis, but also for the opportunity to work closely with her in very exciting research projects and be treated as an equal. Her drive, patience and persistence in combination with her kindness, openness, and friendliness make her a great role model for me. I am especially grateful to Hamsa for all the time she has devoted to make me a better academic writer and presenter.

I also had the unique opportunity to work closely with John Hansman. It was a great pleasure meeting and brainstorming with him. I learned a lot from his out-of-the-box thinking and his insistence on clear and simple ideas. I would like to thank Amedeo Odoni for all the conversations we have had and the continuous mentoring he has offered me since I came to MIT. Furthermore, I would like to thank Eric Feron for his insights, and comments in the later stages of my research. Finally, I would like to thank Richard Jordan and Robert Shumsky for their careful reading of the thesis and thorough comments. I am also grateful to Richard for his feedback during our meetings in the last 5 years.

I feel immense gratitude towards the FAA for funding my research under the PARTNER Center of Excellence, and giving me the opportunity to implement parts of the research in the field. I would also like to acknowledge the cooperation and support of all the individuals who made the field trials at Boston Logan International Airport possible: Brendan Reilly, Deborah James, Pat Hennessy, John Ingaharro, John Melecio, Michael Nelson, Chris Quigley and all the air traffic controllers at the BOS Air Traffic Control Tower Facility; Vincent Cardillo, Flavio Leo, and Robert Lynch at Massport; George Ingram and other airline representatives at the ATA. I am grateful to Tom Reynolds, Masha Ishutkina, James Kuchar, Jim Eggert, Daniel Herring, Rich DeLaura, and Ngaire Underhill of MIT Lincoln Laboratory for their support and help with the ASDE-X and RAPT data. I would also like to thank Alex Nakahara, who provided assistance in computing the fuel burn savings from the field trials, and Regina Clewlow, Alexander Donaldson and Diana Michalek Pfeil, who helped with tower observations before and during the trials. Last but not least, many thanks to Harshad Khadilkar for his great contribution on planning, executing and analyzing the field trials.

I was fortunate to not only be part of the vibrant ICAT lab, but also collaborate on research

papers with many fellow students. The conference papers with Harshad Khadilkar, Hanbong Lee, Nikolas Pyrgiotis, Alexander Donaldson, Regina Clewlow, Lanie Sandberg and the working paper with Alexandre Jacquillat proved to be valuable foundations for the material presented in this thesis. I am really thankful to you guys; it was a real pleasure working with you and learning from you. Thanks also to Claire Cizaire, Lishuai Li, and Diana Michalek Pfeil for all the advice, research and life discussions. Finally, a special thank-you note goes to my statistics and optimization mentor, Varun Ramanujam.

Spending the last five years in graduate school was a truly amazing experience thanks to the company of truly exciting people. In particular, I would like to thank Gerasimos Skaltsas, Nikolas Pyrgiotis, Kostas Bimpikis and Ruben Lobel for the great times we spent in Cambridge and the magnificent traveling experiences we shared in many parts of the world. A special thanks goes to Anurupa Ganguly, for her patience and her empathy in my long and winding road towards the PhD completion.

Finally, I would like to express my deepest appreciation to my parents, Panagiota and Kostas, for their support they have given me while at MIT and their life-long commitment to providing me with the best educational opportunities available. I am also very grateful to my sister, Vasiliki, for her uplifting visits in Cambridge.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The significant increase in air traffic demand in the United States over the past two decades has been accompanied by an increase in flight delays. A detailed analysis of domestic air traffic delays in 2007 by the Joint Economic Committee of the U.S. Senate found that these delays had a $41 billion impact on the nation's economy [65]. The study also estimated that while the major portion of these delays were absorbed when flights were still at their gates, 20% of the delays were incurred as flights were taxiing out to the runway. While some delays, such as those due to extreme weather events, are unavoidable, others can be significantly decreased through better planning and control. Delays caused by an imbalance between available capacity and demand, or delays due to congestion, fall into this category. Operational data shows that in the past decade, more than 15% of National Airspace System (NAS) flight delays have been due to terminal-area volume, or congestion around airports [43]. These delays were incurred when the airports were operating in their optimum configuration, and there were no other impacting conditions. Delays at major airports have also been seen to propagate to large parts of the system [93]. In addition to the increased taxi-out times and delays, airport congestion results in increased fuel burn, emissions, and noise and air quality impacts [21, 82, 85, 107, 129]. The overarching objective of this thesis is the development of approaches for the reduction of the adverse impacts of airport congestion, through the analysis, modeling, and control of the departure process.

1.1 Motivation

Airport surface congestion has several undesirable impacts, the most noticeable of which is the increase in taxi-out times. An analysis of operations in the year 2007 at John F. Kennedy (JFK),

Newark Liberty (EWR) and Philadelphia (PHL) airports showed that they experienced surface congestion, that is, there were more than enough active flights to sustain departure throughput at the airport, 10% to 20% of the time (we will formalize this notion of congestion later in this thesis) [108]. During these periods of congestion aircraft also experienced very high taxi-out times. For example, even under Visual Meteorological Conditions (VMC), the average taxi-out time at JFK was 56 min when the airport was congested, while the unimpeded taxi-out time at JFK was only 16 min in VMC. Similarly, the average taxi-out time at PHL was 38 min when the airport was congested in VMC, while the unimpeded taxi-out time was only 12 min [108].

Figure 1-1 shows the nature of surface congestion during evening operations at PHL[1] [102]. Aircraft taxiing out are depicted in green, and aircraft taxiing in are shown in red. Runway 27L, highlighted with a green arrow, is the departure runway and Runway 27R, highlighted with a red arrow, is the arrival runway. There are approximately 22 aircraft (AC) in the three queues forming near the threshold of the departure runway. We note that it is quite difficult to quantify the exact length of the departure queue, as the 35 aircraft taxiing out seem to be dispersed across multiple taxiways.

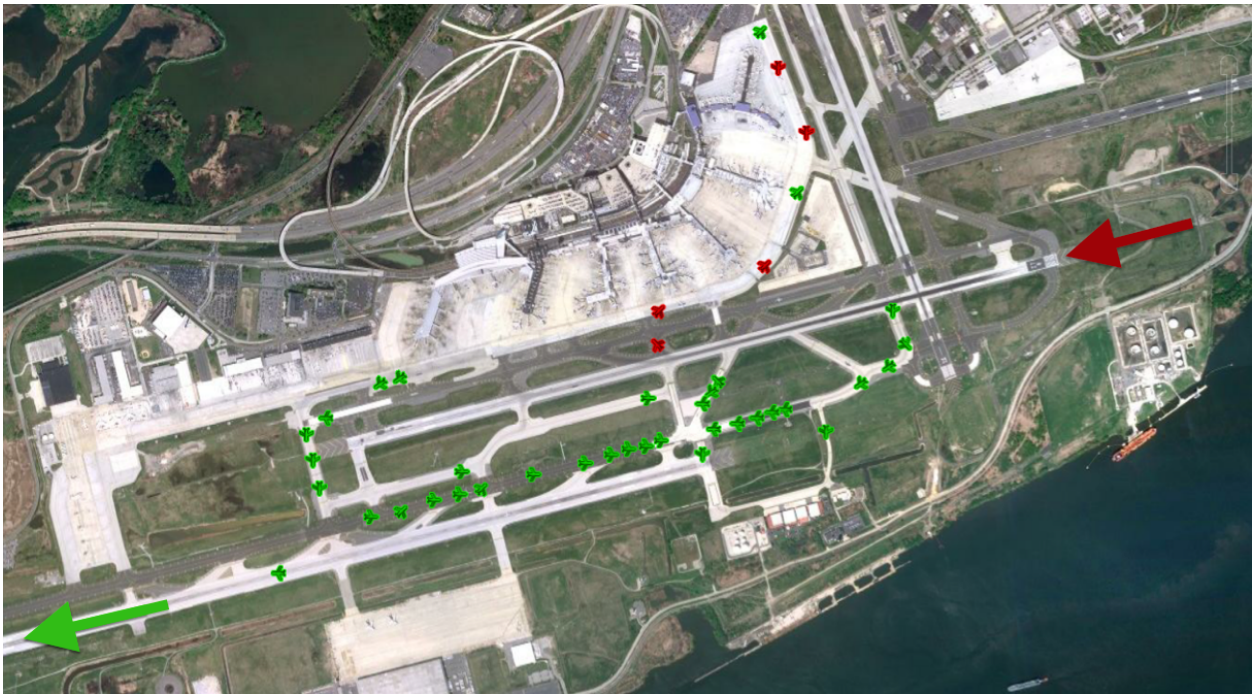


Figure 1-1: Departure queue of Runway 27L at PHL [102].

We use data from the Aviation System Performance Metrics (ASPM) database [38] of the

[1]The visualizations used surface surveillance data from the Airport Surface Detection Equipment – Model X, or ASDE-X system [37].

Federal Aviation Administration (FAA) to further investigate the congestion problem at PHL. In the upper plot of Figure 1-2, we show the average number of pushbacks and the average number of takeoffs (or departures) that was recorded during each 15-minute time window for all days that this runway configuration was in use in 2011. We also show the average departure capacity of this runway configuration of PHL, which is estimated to be 13 AC/15 min, as it will be shown in this thesis. This plot illustrates the imbalance between pushback rate and departure rate. While the average departure throughput is constrained by the capacity of the airport, the demand (pushback rate) can be much higher. The impact of this imbalance on taxi-out times is seen in the lower plot of Figure 1-2, where the average taxi-out times for the flights that pushed back in each 15-minute time window are shown. The figure clearly shows the correlation between excessive pushback rates and large taxi-out times.

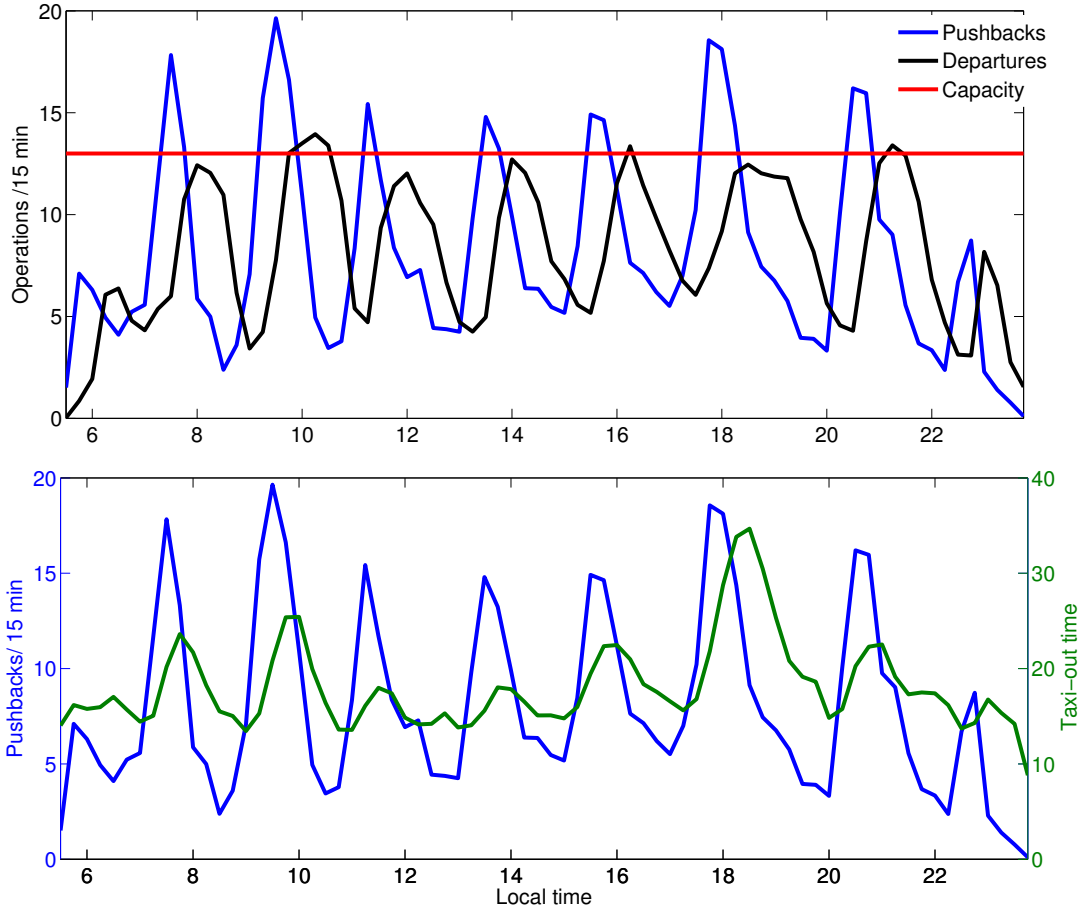


Figure 1-2: Average number of pushbacks, average number of takeoffs and departure capacity per 15 minutes at PHL in 2011 (top); Average number of pushbacks per 15 minutes and average taxi-out times (bottom).

Figure 1-2 summarizes some of the challenges that motivate the work in this thesis. The first question relates to the estimation of the departure capacity of an airport and its distribution under different conditions. It is also desirable to identify opportunities for its improvement, since it is well known that a small increase in capacity can yield large reduction in taxi-out delays [30, 57]. The characterization of the relationship between airport congestion and taxi-out delays is the next objective. Understanding this relationship would help us predict the taxi-out delays that result from the imbalance between demand and capacity, as shown in Figure 1-2. The final challenge is the development of control algorithms that will reduce the impacts of ground congestion, by appropriately regulating pushbacks. Our ultimate goal is to develop tools for the current system that improve the efficiency and predictability of the departure process, and thereby reduce excessive taxi-out times and the associated costs (fuel burn, emissions and controller workload). The proposed approaches do not assume the presence of perfect information and 4D trajectory conformance, but instead try to achieve efficient planning under existing conditions and levels of technology.

1.2 Background and literature review

1.2.1 Characterization of airport capacity

Quantifying the number of arrivals and departures that can be serviced at an airport is important for both strategic planning purposes and tactical air traffic management. Airport capacity is one of the most important inputs for air traffic flow management programs employed in practice by the FAA [34] and proposed in research [12, 121]. It is also the fundamental input for studies aimed at addressing the costs of air traffic delays [94, 125]. Despite the importance of the runway capacity, there is no method available for a simple, consistent and generalizable quantification of airport capacity. Instead, the FAA relies on heuristic estimates of runway capacity based on controller experience and facilities' preferences [40, 41].

Despite the usefulness of the current rules derived by operational experience, the lack of a scientific approach in assessing the capacity of an airport can lead to misconceptions. For example, recent slot limitations at EWR restricted the *total* number of movements in the airport to 81 movements/hour without differentiating between arrivals and departures [36]. However, this thesis will show that the capacity of EWR is best defined as a function of departure and arrival capacity, and not merely their sum. Another common approach is to study congestion pricing under the assumption that all users impose the same external marginal cost on other users, all else being

equal [35]. However, this thesis will show that the external marginal cost of a specific type of user can differ significantly depending on the specifics of the airport. Such differences must be identified before being used for the design of policy instruments.

Runway capacity estimation

It has been shown that the most fundamental quantity that characterizes the performance of the airport is its runway capacity, since the main throughput bottleneck at an airport is the runway system [61]. Along the same lines, runway capacity has been shown that it is the predominant cause of the most extreme instances of delays [8]. Thus, the capacity of an airport is in most cases synonymous of the capacity of the runway system.

The most influential work in the area of the empirical measurement of the capacity of an airport is that of Gilbo [49], in which he proposed a quasi-statistical procedure for estimating the capacity envelope of a single airport, for a given runway configuration. 15-minute arrival and departure counts were used to estimate the capacity envelope as the convex hull of the scatter of the counts after correcting for outliers. Frequency-based filtering was employed to eliminate outliers. Along similar lines, Ramanujam and Balakrishnan proposed a systematic statistical approach for estimating intra- and inter-airport capacity envelopes from observed data by applying quantile regression [97]. In the above-referenced papers, the capacity envelope was estimated using statistically significant *maximum* counts of the movements. These estimated counts are probably achievable only under certain circumstances, such as a favorable fleet mix, or favorable operations sequencing. Thus, the estimated values do not represent expected number of movements, but rather the upper bounds for the number of movements. In addition, these estimates of airport capacity do not reveal its variance, caused by factors such as fleet mix, weather and airspace restrictions.

Theoretical methods have also been employed for measuring the capacity of the runway system [30]. However, they heavily rely on stylized models and assumptions which tend to vary among different airports. One such example is that wake vortex separation is the prevalent determining factor for the departure capacity. It turns out, that in many cases, airspace constraints are more important than wake vortex separation constraints, as it will be shown in this work. Thus, in certain cases, stylized models can over-estimate the capacity. Theoretical models provide a very good benchmark for identifying opportunities for improvement by comparing their estimates to the empirical ones [64].

New York Metroplex capacity estimation

A critical component of the NAS is the airport system of New York, comprising John F. Kennedy International Airport (JFK), Newark Liberty International Airport (EWR), LaGuardia Airport (LGA), and smaller regional airports. According to the Government Accountability Office (GAO), 41% of total delayed departures and 47% of the total delay in the NAS in 2009 were attributed to the three major US airports (JFK, EWR and LGA) [124] . In response, the Regional Plan Association laid out a comprehensive plan with options for adding capacity and reducing delays in New York airspace [130]. Many researchers have recommended several solutions for improving the imbalance between demand and capacity in the three New York airports [9, 93, 126]. Despite the great level of interest that the major New York airports have attracted from the operations, research and policy communities, few systematic studies have estimated the capacity of the individual airports as well as the Metroplex.

1.2.2 Modeling of departure operations

Prior work on the modeling of the departure process at airports can be broadly classified into three groups. The first group focuses on computing runway-related delays under dynamic and stochastic conditions [71, 81, 94]. This runway-centric approach is justified by the observation that the main throughput bottleneck at an airport is the runway system [61]. This approach views the runway complex of an airport as a queuing system whose customers are aircraft that need to land or takeoff. The models are then used to predict the expected system behavior, and their results are typically most useful for long-term planning (for example, estimating the expected reduction in delays from the construction of a new runway), or for estimating the network propagation of local disruptions, such as thunderstorms. However, their level of abstraction is too high for studying taxi-out reduction algorithms, or for predicting taxi-out times for individual flights. For this reason, these models are typically used to estimate "treatment effects" in the entire NAS. The models are used to calculate a baseline and then estimate the effect of network modifications, disruptions to the system, scenarios of traffic demand, future technologies, or next-generation operations by comparing the results of the test scenario to the baseline scenario. However, the simplifying abstraction of modeling the airport as a runway server may influence these calculations. Recent research has also shown that airspace constraints and downstream restrictions can also lead to significant delays in the NAS [31, 32, 100]. In addition, these models use a Poisson process for modeling the aircraft

service requests at the runway server, which may be an appropriate approximation for landings [128], but has not been validated for departures. Recently, it was shown that simulating the service requests with a less random process, that is, one with smaller support and variation around the expected value, predicts both expected delays and their variability in congested airports [62].

The second category of prior research focused on predicting taxi-out times. Shumsky developed a model to predict taxi times using a variety of explanatory variables such as the airline, the departure runway and departure demand [104, 105]. He also developed a queuing model for the runway service process. However, the queuing model was based on cumulative behavior and did not reflect the stochastic nature of the process [104]. Idris et al. analyzed the main causal factors that affect taxi times and based on this analysis, they developed a statistical regression model to predict taxi times [58], as part of Departure Planning Project [47]. That work did not explicitly model the runway service process, and so could not link the excessive taxi times with the capacity constraints. It could therefore not be used for strategic surface flow management applications such as the one considered in this thesis, where we like to consider gate-to-runway traffic states, and determine how surface queues can be managed in order to reduce taxi-out times.

In contrast, Pujet et al. extended some these notions to predict taxi times using a simple stochastic queuing model [92]. They assumed that an aircraft needs a certain (fixed) amount of time, defined to be the *travel time*, to reach the departure runways. In their model, upon reaching the departure runways, aircraft line up in the *runway queue*, where they get served by the runway server according to a probabilistic service process. Pujet et al. estimated the travel time for each flight based on several casual factors and also modeled the probabilistic service process. Given a pushback schedule, the estimated taxi-out time is the sum of travel time and the wait time for service (takeoff) at the runway queue. In earlier work, we provided with a more complete queuing model of the departure process, by using better unimpeded taxi-out time estimations, new models for the runway server(s) and the ramp and taxiway delays [107]. However, none of the above surface models account for the impact of arrivals nor the air traffic flow management programs, although it has been shown that they have a high impact on taxi-out times [59]. In addition, these models simulate the runway service process using random number generators and thus the output is just a random sample path. Multiple runs are necessary in order to obtain statistically significant estimates.

Based on the progress of work on surface models, some queuing models of the NAS have tried to incorporate a ground component instead of just modeling the airport as a runway server [24, 79, 127].

In particular, the Detailed Policy Assessment Tool (DPAT) built on this approach by adding a stochastic component for the taxiway related delay [127]. This resulted in the first system-wide model to include delays that are not related to the runway queuing time. However, the runway module of DPAT is a deterministic one and the taxiway delay is a user-input. It is not shown how it can be derived, or modeled to reflect the actual taxiway-related delays. In LMINET [79], taxiway delays are modeled by an M/M/1 queue. The assumption that the taxiway system can be modeled as a single server with exponential service rate is debatable. Moreover, the runway related delay model in LMINET is not validated by empirical data. In summary, both DPAT and LMINET suffer from oversimplifying assumptions in the runway-and-taxiway-related delay estimation, representation and prediction. It must be noted that the airport modules of these models have not been fully validated. We hypothesize that such models will not be able to predict taxi out times accurately and as a result, delays may be estimated inaccurately. This inaccuracy may also result in errors in the estimation of the network-wide effects of airport delays.

Finally, a third body of work involves the microscopic modeling of all airport components, such as the Airport and Airspace Simulation Model (SIMMOD) and the Total Airspace and Airport Modeler (TAAM) [87]. These tools model the layout of an airport, the operating rules for every aircraft type, and the dynamics of every gate, taxiway and runway with high fidelity. They need extensive adaptation of both the airport layout and the traffic scenarios so as to generate statistically significant results. It is therefore difficult to use them to perform a probabilistic analysis of the departure process or to test new strategies, such as control of the departure process for emissions reduction, because this would require simulating them over long periods of time [87].

Unimpeded taxi-out estimation

There has been relatively little prior analysis of unimpeded taxi times at airports. Unimpeded taxi-out times are very crucial in understanding the performance of the airports [108], as they need to be subtracted from taxi-out times in order to estimate taxi-out delays. They are also crucial for modeling the departure process.

The FAA defines the unimpeded taxi-out time as *the taxi-out time under optimal operating conditions, when neither congestion, weather nor other factors delay the aircraft during its movement from gate to takeoff* [88]. The unimpeded taxi-out time is redefined in terms of available data as *the taxi-out time when the departure queue is equal to one and the arrival queue is equal to zero.* A linear regression of the observed taxi-out times with the observed departure and arrival queues is

then conducted, and the unimpeded taxi-out time is estimated from the linear regression equation by setting the departure queue equal to 1 and arrival queue equal to 0 [46].

Idris et al. [58] observed that (1) there is poor correlation of the taxi-out times with arriving traffic, and (2) the taxi-out time of a flight is more strongly correlated with its takeoff queue than the number of departing aircraft on the ground. The exact dependence of the taxi-out time of an aircraft with the takeoff queue, and how this dependence can be used for estimating the unimpeded taxi-out time, both remain unanswered questions.

Nevertheless, researchers have proposed solutions for better estimating the unimpeded taxi-out time [20, 92]. However, all of these methods are heuristic improvements of the FAA method. For example, Pujet [92] extended the method of FAA to consider a range of values 0-2 of the departure queue. He also derived distributions instead of point estimates. We used a method based on the concept of the takeoff queue, but we assumed normally distributed errors around the estimated value [107]. However, it is known that unimpeded taxi-out times are not normally distributed in practice [70, 92].

We also note that the recently available ground surveillance (ASDE-X) data enabled the extraction of unimpeded taxi-out times by directly estimating the time an aircraft needs to cross a taxiway link conditioned on this link being free of other traffic [70, 77]. ASDE-X coverage is poor in the ramp area, and thus it renders ASDE-X data use unreliable for extracting the unimpeded travel time from the gate to runway. In addition, we would like to develop a more general method which does not need to rely on readily-available ASDE-X data capability.

1.2.3 Optimization of the departure process

Taxi-out time optimization

In most recent research, the control of the airport departure process is formulated as a surface traffic optimization problem [75]. In these formulations, the airport taxiway system is modeled as a network of links and nodes. In this node-link model, nodes represent significant control points on the airport surface, such as gate locations, runway entry and exit points, intersections of taxiways, and holding spots for clearance. The taxiway segments between two points on the surface are represented as links between the respective nodes. Flights get routed through the network from a starting to an ending point. If the flight is a departure, the starting point is its gate and the ending point is the assigned departure runway. The surface traffic optimization problem is formulated as

a scheduling or routing problem for the flights through the network. The objective of this problem is to minimize taxi delay given a runway schedule. The optimization is performed by choosing between alternative routes through the network, by scheduling the passage times on the nodes, or a combination of both.

This optimization problem is typically solved as a Mixed Integer Linear Programming (MILP) problem. Smeltink et al. were the first ones to develop a MILP model to determine the movement of taxiing aircraft and meet basic safety and operational constraints for the Amsterdam airport [115]. The model, however, had long computation times and missed some constraints such as the runway occupancy time. Rathinam et al. improved Smeltink et al.'s model and applied their approach to simulations at Dallas Fort Worth (DFW) airport [98]. They incorporated more operational constraints, such as the aircraft type for wake vortex separation requirements. However, the model, tested with empirical data, showed very long computation time for high density traffic. Balakrishnan and Jung proposed an integer programming formulation for optimizing surface operations at DFW by adapting the Bertsimas and Stock-Patterson formulation [12] for the Traffic Flow Management Problem [7]. Through simulations with actual DFW airport data, they evaluated two strategies for improving the taxi times; controlled pushback and taxiway reroutes. This model improved the formulation for the surface operations optimization and its computational performance, but did not account for several operational restrictions such as overtaking constraints and collision avoidance at intersections. This formulation was further refined by Lee and Balakrishnan who considered additional operational constraints [75].

This problem has, in its general form, been shown to be NP-hard [12, 115]. Practical implementations consider scheduling a small number of operations, typically 20-30 flights at a time. The problem of scheduling a day of traffic at an airport is recast as a rolling-horizon problem, with a typical horizon being 15 minutes. In other words, the solutions are open-loop policies subject to periodic reoptimization. It is not clear how suboptimal the rolling-horizon solution is with respect to the global solution. The robustness of the optimal solutions has not been extensively investigated. In addition, perfect information of the location and the intent of all aircraft on the surface, deterministically known schedules, and, in many cases, ability of the decision maker to instantaneously set and change aircraft states and speeds are assumed. It is not analyzed how well these algorithms would perform in a dynamic environment in the presence of uncertainties, or unexpected events, such as an aircraft taking too long to pushback, a mechanical problem, a datalink failure, a safety incident, a temporary runway closure, etc. In related work, we have shown that applying the

optimized schedule without enforcing pilot conformance to the specified arrival time at the network nodes can end up being suboptimal even compared to current operations [76].

There has been little work on the stochastic optimization of the taxiing operations. One possible reason for this is that the variability of the underlying processes is of the same order of magnitude as the quantity to be optimized. For example, if the duration of the pushback process is uniformly distributed between 2 and 5 minutes, and the optimized taxi-times are in the order of magnitude of 15 minutes, it may be very hard, or not really useful to formulate the problem as a stochastic routing problem through a network where inter-arrival times at successive nodes are as short as 30 seconds. Another challenge is that the computation requirements of the stochastic formulation of this problem can be prohibitive.

There have also been a few alternative, non-MILP formulations to this routing problem. For example, Gotteland et al. use genetic algorithms to choose between alternative taxi routes [50], and Trani et al. use Time-Dependent Shortest Path techniques [122]. However, little has been shown about the optimality, performance or robustness of these approaches.

Optimization of runway operations

A problem closely linked to the control of taxi times is that of runway operations scheduling. Clearly, a more favorable runway schedule yields higher throughput and consequently shorter taxi times for the aircraft in the sequence. In contrast, bad sequencing at the runway will mean higher taxi times for everyone and could diminish any benefits of taxi-out time optimization or control.

Nowadays, runway scheduling at all major airports is mostly done on first-come-first-serve (FCFS) basis. Air traffic controllers consider heuristic deviations from the FCFS discipline in order to increase departure throughput[2]. For example, in Figure 1-1, we observe three queues feeding the departure runway. Air traffic controllers at PHL feed aircraft in different queues so as to increase opportunities for dispersal headings and improve the departure throughput of the airport [99].

In the research literature, the problem of scheduling runway operations has been particularly studied for the case of landings. It has been shown that the problem of optimal runway operations sequencing is in its general form NP-hard [10]. However, for practical applications, one only needs to consider deviations from the FCFS discipline only within a specified maximum number of positions. Such a policy is called constrained position shifting (CPS). The CPS problem has complexity that scales linearly with the number of aircraft [6]. The CPS framework has also been extended for

[2]Interviews with air traffic controllers from the BOS and PHL Air Traffic Control Tower facility.

robust runways operations planning allowing for uncertainty in the estimated time of arrival (ETA) of an aircraft [22].

The problem of scheduling and sequencing departures at a runway exhibits significant differences from the problem of landings. In the case of the landings, all aircraft are airborne and have an ETA at the runway. This ETA can be assumed to be deterministic or probabilistic and the scheduling and sequencing can be efficiently formulated as a dynamic programming recursion, as discussed previously [6, 22]. However, in the case of departures, the ETA of aircraft at the runway is also a decision variable. If a taxiway planner is available, it will attempt to minimize taxi times by delaying flights at the gate as long as possible. Furthermore, estimating the time of arrival at the runway of a departing aircraft after clearing it for pushback is a much harder problem. As we shall see, the inherent uncertainties are much larger. These systematic uncertainties are mitigated today by ensuring that there is a large pool of aircraft available for takeoff. This is at odds with the taxiway planner, which would attempt to minimize the number of aircraft on the surface. Finally, departures scheduling has occasionally to be traded with other events, like runway crossings, or arrivals scheduling. These additional constraints are absent from the arrivals scheduling problem, because arrivals are by default given priorities over these events.

There is relatively little work in the literature addressing the peculiarities of the departure scheduling problem. Anagnostakis et al. addressed some of them, by incorporating constraints such as runway crossings, minutes in trail, miles in trail and Expect Departure Clearance Time (EDCT) requirements in the departures scheduling problem [1, 2, 3]. In particular Anagnostakis and Clarke proposed a two stage solution approach: In the first stage, the optimal weight class sequencing is heuristically determined. In the second stage, individual aircraft are assigned to available weight class slots. The first stage maximizes throughput and the second stage minimizes individual aircraft delay [1, 2]. Balakrishnan and Chandran proposed algorithms for departure scheduling in the presence of position shift constraints [5]. However, the suggested heuristics and algorithms were deterministic planning tools. More recently, a stochastic departure runway planning tool was proposed by Solveling et al. [117] extending the two-stage deterministic approach proposed by Anagnostakis and Clarke. This was the first stochastic planning tool tailored to the problem of departures scheduling addressing the uncertainty of the pushback process, the taxiing process and the time of arrival of landing aircraft. Nonetheless, it did not incorporate restrictions from air traffic flow management programs. Additionally, the proposed algorithm had very high computational cost, and an approximate solution could be found only for small sequences of flights, such as eight

flights. For this reason, a rolling horizon approach was implemented.

Integration of taxiway and runway operations

While the problems of taxiway and runway operations optimization are clearly coupled, most works have treated them separately. On the optimization front, there is one previous example of combining a taxiway and a runway planner. As part of the NASA Airportal project [23], the taxiway planner of Lee and Balakrishnan [75] was sequentially combined with the runway planner of Solveling et al. [117]. The runway planner produced an optimal runway schedule, and the taxiway planner minimized the taxi times that achieved the prescribed runway schedule. Preliminary results suggest that the taxiway planner provided the most benefit in terms of reducing taxi-out time, that the runway planner detracted from the positive effects of the taxiway planner, and that a FCFS runway scheduler is sufficient for the traffic scenarios examined [52].

On the taxi-time control front, runway operations are typically viewed as a constraint that the control algorithms must not violate: The taxi times must be controlled with sufficient robustness so that there is always sufficient demand at the runway threshold. With the exception of the work of Burgain [15], authors do not explicitly quantify what this sufficient demand is, but rather give some conditions which are supposed to guarantee persistent demand at the runway. The problem of synergistically reducing both the taxiway and runway inefficiencies is not examined.

1.2.4 Control of departure processes

In the realm of air traffic flow management, there is a class of alternative models to the ones that do aircraft-based optimization. These models, called Eulerian models, are concerned with the optimal control of the flow of aircraft, rather than the trajectory of individual aircraft. Eulerian models only deal with aircraft counts in specific control volumes of airspace rather individual aircraft trajectories, and are more tractable for the purpose of control [74, 101]. However, these dynamic control approaches have not been applied to the aircraft flow on the surface of the airport.

An airport congestion control strategy in its simplest form would be a state-dependent pushback policy aiming at reducing congestion on the ground. One such approach is the *N-Control* strategy. N-Control is one implementation of the virtual queue concept described in the Departure Planner [47] and variants of it have been extensively studied [16, 19, 20, 92]. The main idea behind N-Control is an observation of the performance of the departure throughput of US airports: As more aircraft pushback from their gates onto the taxiway system, the throughput of the departure

runway initially increases because more aircraft are available in the departure queue. However, as this number, denoted N, exceeds a threshold, the departure runway capacity becomes the limiting factor, and there is no additional increase in throughput. We denote this threshold as N^*. The dependence of the departure throughput on the number of aircraft taxiing out is illustrated in Figure 1-3 using ASPM data from 2011 for runway configuration (VMC; 31 | 4) of LGA. Beyond the threshold N^*, any additional aircraft that pushback simply incur taxi-out delays without increasing the airport throughput [107].

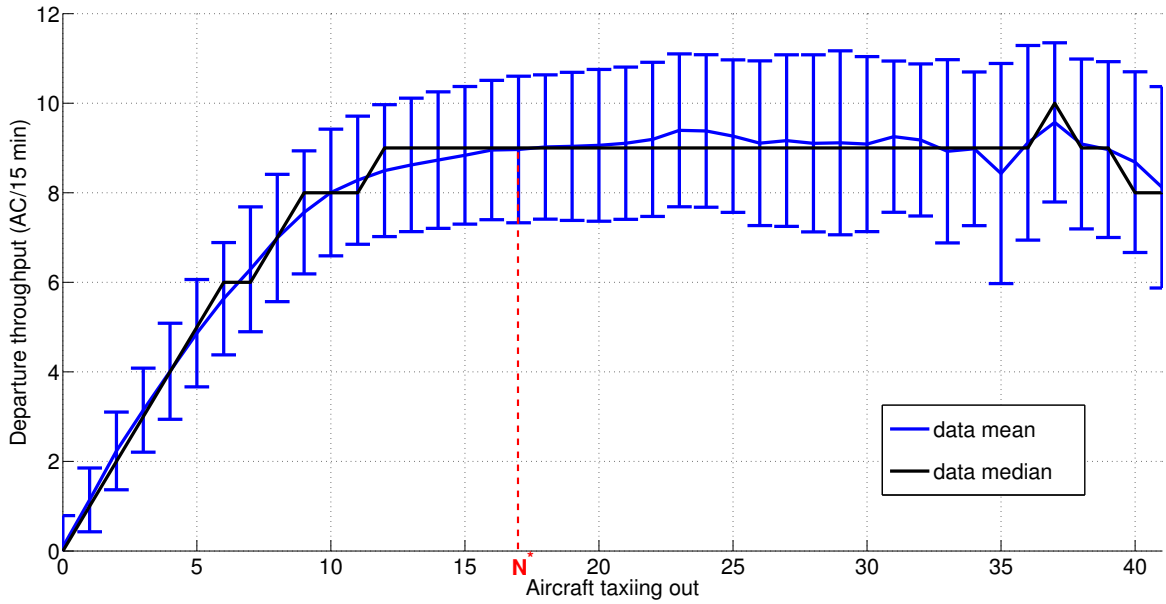


Figure 1-3: Departure throughput as a function of the number of aircraft taxiing out, for the (VMC; 31 | 4) configuration at LGA

The policy is effectively a simple threshold heuristic: If the total number of departing aircraft on the ground exceeds a certain threshold, N_{ctrl}, where $N_{ctrl} \geq N^*$, stop dispatching aircraft requesting pushback until the number of aircraft on the ground drops below the threshold. While the choice of N_{ctrl} must be large enough to maintain runway utilization, too large a value will be overly conservative, and result in a loss of benefits from the control strategy. Such a policy aims at preventing excessive congestion and is already heuristically in use by air traffic controllers [25] during excessively congested situations.

The N-Control policy is also closely related to the *constant work-in-process* (CONWIP) policy in manufacturing systems. The main benefits of CONWIP systems are their simplicity, implementability and controllability [118]. This approach presents an efficient way to control congestion

by accepting an adjustable risk of capacity loss. The N-Control strategy has been shown through several simulation studies to have similar properties.

More complex policies which attempt to attain some optimization objective have not been considered for surface traffic until recently. In 2009, Burgain et al. used more advanced modeling and optimization tools for the characterization of optimal pushback policies. More specifically, they modeled the airport surface with a state space model, and characterized optimal pushback policies as a function of the state of the system, and not just the total number of aircraft on the ground [17]. The optimal policies considered were full-state feedback policies, which faced some implementation issues [15]. This control protocol was a generalization of the N-Control strategy, in which, at each minute, the state of the surface was mapped to an *on-off* input signal.

All the above policies (N-Control, CONWIP systems, and Burgain et al.'s refinements of N-Control) can be classified as token-based, or surplus-based policies [48]. In these approaches, every state transition generates a token, an action or a signal, which is applied at the input to the system (the pushback process). Equivalently, every state transition translates to a new surplus level (or lack thereof) at different buffers of the system, which implies a different flow of input into the system. More general approaches can be found in the literature on the dynamic control of queuing systems.

There has been much prior research on the optimal control of a variety of queuing systems, considering different decision variables and control objectives [29, 78, 80, 119, 120]. In the modeling framework of Low [80] and Lippman [78] , the state of the system is the number of customers in the system. A holding cost h_i is incurred per unit time that the system is in state i and a reward (or entrance fee) p_λ is received (paid) whenever a customer enters the system at a point in time when the arrival rate is λ. h_i is assumed to be convex and non-decreasing. Then, it can be shown that the optimal λ is a non-decreasing function of the state i. This formulation can be easily modified so as to yield optimal policies with respect to performance measures such as throughput, congestion, or a combination of the two.

Several challenges remain when attempting to apply results from queuing, manufacturing and inventory control in the context of controlling the departure process. Firstly, on-off or event-driven control policies for controlling the pushback process are difficult to implement in practice. Both the air traffic controllers and the airlines would prefer a state-dependent pushback rate that would be valid for a predefined time period, after which it would be updated. Air traffic controllers prefer such periodically updated pushback rate recommendations for workload and procedural reasons,

and airlines prefer them because of their predictability, which is essential for planning ground crews. Secondly, the control input is applied at the gates during pushback, whereas the main bottleneck is the runway. The control strategy cannot be applied directly at the runway queue, but instead has to accommodate stochastic taxi-out times between the gate and the runway. Factors that contribute to the stochasticity of taxi-out times include the pushback process, flight checklists, communication delays, and variable taxi speeds.

For all of the above reasons, the N-Control strategy has not been applied as such in the airport context. Researchers have developed several heuristic modifications of it for field-testing customized for different airport environments and requirements. Characteristic examples of these efforts are the metering of departures at New York JFK airport by PASSUR Aerospace, Inc. [86], the field evaluation of the Collaborative Departure Queue Management concept at Memphis (MEM) airport [14], the human-in-the-loop simulations of the Spot and Runway Departure Advisor (SARDA) concept at DFW airport [66] and the trials of the Departure Manager (DMAN) concept [13] in Athens International airport (ATH) [103]. During summer of 2010, we also developed and successfully tested such a heuristic, the Pushback Rate Control protocol (henceforth referred to as PRC_v1.0) [110]. However, none of these efforts, which are essentially different variants of N-Control [85], have explicitly estimated the stochasticity of the underlying processes, and developed algorithms that explicitly attain certain objectives.

Finally, when applying results from queuing theory to congestion control, another critical issue is that the service times are not exponentially distributed [55, 107, 113]. Researchers have considered modeling them as time-dependent Erlang distributions [55, 94, 113], time-dependent deterministic distributions [73, 89], binomial [91], or multinomial distributions [15, 107]. However, to the best of our knowledge, empirical service times have not been extracted to date from operational data.

1.3 Main contributions of this thesis

This research takes important steps towards better understanding, modeling and tactically managing the airport departure process. The main contributions of the research are:

1. The development of a new method for the characterization and parametric estimation of airport capacity.

2. The analysis of the capacity of the New York Metroplex through a detailed case study.

3. The development of a new analytical model for predicting unimpeded taxi-out times, taxi-out times and departure queuing delays.

4. The algorithmic development of a class of airport surface congestion control algorithms based on dynamic programing, and their evaluation through simulations.

5. The detailed evaluation of the field-testing of an airport surface congestion control algorithm at Boston Logan International Airport (BOS) in summer 2011.

The contributions are briefly described in the following sections.

Characterization and estimation of runway and airport capacity

We introduce a new concept for characterizing the capacity of an airport which we name *operational throughput envelope*. This method measures the average departure throughput under persistent departure demand as a function of the average arrival throughput. We show how the conditions for persistent demand can be guaranteed and compare our results to other capacity estimation methods. We apply the methodology to the estimation of the operational throughput envelope for major airports with different layouts, such as Boston Logan International Airport (BOS), Newark Liberty International Airport (EWR), LaGuardia Airport (LGA), John F. Kennedy International Airport (JFK), Charlotte Douglas International Airport (CLT), and Philadelphia International Airport (PHL). We also apply it for measuring the capacity of runway 17R of Dallas Fort Worth Airport (DFW). In all these cases, we provide operational throughput envelopes that can be used for better understanding the tradeoffs between arrival and departure operations at each airport, for strategically planning of operations and for identifying opportunities for improvement.

We extend the *operational throughput envelope* method to estimate the dependence of the departure throughput on parameters other than the arrival throughput. We show with several examples how the explanatory variables can be chosen for different airport environments and the insights provided by the analyses. For the case of BOS, we show that the ratio of propeller-driven aircraft in the fleet mix is a more significant explanatory variable of the departure throughput, than is the arrival throughput. For EWR, we show that the departure throughput is relatively inelastic to the fleet mix of the departing aircraft.

We also introduce a recently available data-source, the Route Availability Planning Tool (RAPT), and explain how it can be used for estimating the throughput of LGA during convective weather. In particular, we measure the impact of available airspace capacity on the operational throughput

of the airport. To the best of our knowledge, this is the first study to use route availability information, and not simply local weather conditions, to characterize the available capacity of an airport. Portions of this work appeared in [110, 112].

Capacity analysis of the New York Metroplex

We then extend the parametric estimation methodology to study interactions among different airports, in particular the three major airports of New York Metroplex, JFK, EWR and LGA. We show that operations at one airport do not adversely impact operations at another airport of the Metroplex. We attribute this due to the current airspace design and operational procedures, which keep operations at the three airports segregated. We conclude by characterizing the capacity of the Metroplex under different configurations and conditions. Portions of this work appeared in [112].

Departure process modeling

We introduce a method that estimates the taxi-out times as a function of the traffic in the airport. From this, we extract an empirical distribution for the unimpeded taxi-out times. We then estimate the parameters of the runway service process using operational data and derive probability distributions for the service times of the process and the capacity of the airport. We then develop a new stochastic and dynamic queuing model for modeling the queuing delay at the departure runway(s), based on the transient analysis of $D(t)/E_k(t)/1$ queuing systems. The model takes as input the pushback schedule, derives the expected runway schedule and yields the expected takeoff times. It provides estimates of the average taxi-out times, the average queuing delays and their variance for each flight, the congestion state of the airport, the load of the departure queue(s) and the departure throughput.

To the best of our knowledge, this model is the first analytical model to provide estimates for the variance of the queuing delays. The analytical model yields very accurate predictions for the airports analyzed, EWR, PHL and CLT. We also show how Monte Carlo simulations can be run using the proposed framework to derive distributions for the taxi-out times. We demonstrate the superiority of the stochastic model compared to a deterministic one and show that there is insignificant benefit added from running detailed Monte Carlo simulations in lieu of the analytical model for estimating expected delays. Portions of this work are described in [26, 76, 113].

Dynamic control of airport departures

Here, we take several important steps with regard to the control of the airport departure process. Using the findings from the previous sections, we address three main challenges of controlling the departure process: the random delay between control actuation (at the gate) and the server being controlled (the runway), the stochasticity of the runway service process and the need to develop control strategies that can be implemented in practice by human air traffic controllers. We propose control algorithms that predict the departure throughput in the next time window, and recommend a rate at which to release pushbacks from the gate in order to control congestion and the associated taxi-out times, fuel-burn and emissions. We name such algorithms *Pushback Rate Control (PRC)* strategies. We derive two control algorithms using dynamic programing and approximate dynamic programming. To this end, we model the runway system as a semi-Markov process, and we introduce two new queuing models for modeling the *controlled* departure process, the $(M(t)|R)/E_k/1$ and $(M(t)|R)/D_s/1$ model. These models contribute to the literature of transient queuing systems analysis, as they assume a deterministically known number of aircraft, R, randomly reaching the departure runway during a predetermined time-window. The proposed approach is trained with ASPM data, simulated at PHL airport and compared to other control policies. Portions of this work appeared in [109].

Testing of a Pushback Rate Control Strategy

We also train the controlled queuing model using ASDE-X data, and visual observations from BOS and derive optimal pushback strategies for this particular airport. We describe the field-testing of one PRC strategy at BOS during summer 2011. In addition to providing optimal pushback rates in a real time operational environment, we design and implement a decision support tool that is used by air traffic controllers during the field testing. Finally, we simulate the "what-if" scenario and provide an extensive evaluation of the congestion management scheme, the takeoff rate prediction accuracy and the implementation. To the best of our knowledge, this work offers the only detailed evaluation of congestion management techniques. Portions of this work were published in [110, 111].

1.4 Organization of the thesis

The thesis is organized as follows. Chapters 2 and 3 deal with the analysis of airport departure capacity. In Chapter 2, we describe the methodology to derive the average throughput envelope

and the parametric estimation of the departure capacity. We demonstrate the approach by using the example of BOS. In Chapter 3, we perform a capacity analysis for the New York Metroplex, including the average throughput estimation for EWR, LGA and JFK, and the parametric analysis of EWR departure capacity and LGA capacity. We also study the interactions among the three airports, and characterize the capacity of the Metroplex. In Chapter 4, we develop an analytical model of the departure process, and test its performance in predicting delays at EWR. In Chapter 5, we develop an algorithm for the dynamic control of the departure process. We also describe the field-testing of a Pushback Rate Control strategy at BOS, and give a detailed evaluation of the field-test. In Chapter 6, we summarize our main findings and discuss future directions of this research. The appendices provide additional examples and case studies, including analyses of single runway performance at PHL and DFW, prediction of delays at PHL, and assessment of infrastructure improvements at Charlotte International Airport (CLT).

Chapter 2

Characterization and Estimation of Airport Capacity

The capacity of an airport is typically measured by the average number of movements that can be performed on the runway system in the presence of continuous demand [30]. The runway system is shared by departures and arrivals, and therefore the capacity of a runway system is represented by the number of operations achieved at specific arrival/departure mix ratios, also known as the capacity envelope [30].

In this chapter, we propose a new methodology for estimating and representing the capacity envelope of a runway system. We also show how this methodology can be used to measure and visualize the departure capacity conditioned on several explanatory variables, such as fleet mix, arrival throughput etc. We first illustrate the steps involved by applying the proposed framework to ASPM data [38] from Boston Logan International Airport (BOS) for one major runway configuration in 2007. Subsequently, we show results for all major runway configurations at BOS, validate the results with ASPM data from 2008, and compare them to those obtained with a different data-source (ASDE-X).

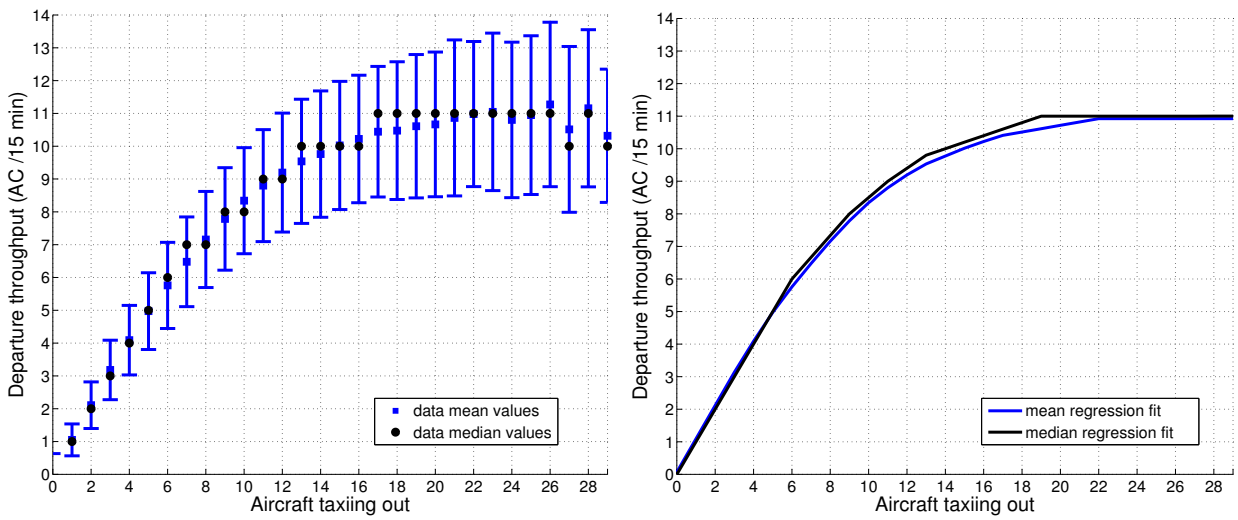
All estimation problems are formulated as convex optimization problems, and solved using the CVX MATLAB-based modeling software [51].

2.1 Average departure throughput

The first step in measuring the capacity of an airport is the estimation of the average departure throughput. For this, we represent the departure throughput as a function of departure demand.

The departure demand, $N(t)$, at some time t (represented in 1-min increments), is defined as the number of aircraft taxiing out during that time interval. In other words, it is the number of aircraft that have pushed back from their gates, but have not taken off yet. The departure throughput during a 15-minute period starting at time t, denoted $T(t)$, is measured as the number of aircraft that take off during the 15-minute interval $[t, t + 15)$ min.

This representation, introduced by Shumsky [104] and used by Pujet [91], yields plots such as those shown in Figure 2-1a for the most frequently used runway configuration 22L, 27 | 22L, 22R at BOS in 2007 under Visual Meteorological Conditions (VMC). For the entire year, we have 121,414 data points of the form (N, T). In Figure 2-1a, we plot the mean and median departure throughput for each value of the departure demand, N. The error bars depict the standard deviation of the departure throughput at each value of N. As the number of aircraft taxiing out increases, the



(a) Measurements of the departure throughput as function of the number of aircraft taxiing out.

(b) Regression of the departure throughput as function of the number of aircraft taxiing out.

Figure 2-1: BOS throughput in configuration (VMC; 22L, 27 | 22L, 22R).

departure throughput initially increases, but then saturates once the demand exceeds a critical value N^*. From Figure 2-1a, we can infer that N^* is around 20 aircraft and that the average departure throughput in saturation is around 11 AC /15 min. Similar behavior was observed by Shumsky [104], who used Airline Service Quality Performance (ASQP) [44] and Enhanced Traffic Management System (ETMS) [45] data from year 1991, and Pujet [91], who used ASQP data from year 1995. However, the parameters N^* and saturation throughput, calculated by Pujet and Shumsky, are significantly different from those estimated in this work. These differences can be

explained by the observation that their data included only 53% and 65% of flights respectively. In the remainder of this thesis, we refer to plots such as those of Figure 2-1 as *saturation plots*.

We formalize the estimation of the mean and median throughput rate as functions of the departure demand by formulating the estimation problems as regression problems. The functions f_{mean} and f_{med} to be fitted have to adhere to the physics of the system, namely, that:

- Departure throughput is a monotonically non-decreasing function of departure demand.
- Departure throughput is a concave function of departure demand.

We briefly explain why the departure throughput is expected to be a concave function of departure demand. Suppose that increasing the demand from $(x-1)$ AC to x AC implies an increase of the departure throughput by y AC/15 min, where $y \leq 1$. This increase can be interpreted as the probability that the xth AC departs in the next 15 min. When increasing the departure demand from x AC to $(x+1)$ AC, suppose that the departure throughput increases by $z > y$. If there is enough capacity for the $(x+1)^{\text{th}}$ AC to depart with probability z, the probability of the xth AC departing is at least z, because if the xth AC does not depart, the $(x+1)^{\text{th}}$ AC can clearly not depart. This implies that $z \leq y$, which is a contradiction. Thus, the departure throughput is a concave function of departure demand.

We also note that there are cases in which very high departure demand can be associated with decreasing throughput. An example of such a case is shown in Appendix G for PHL. One potential reason for a decreasing trend in the empirical data is that very high values of traffic can lead to gridlocks on the surface and loss of runway capacity. For example, all taxiways surrounding the departure runway may be occupied by aircraft, and an aircraft may need to be routed through the departure runway.[1] In addition, as we shall see in the subsequent sections, there are are other variables that explain the variation in the departure throughput apart from the departure demand. For example, downstream restrictions can lead to decreased capacity, which can lead to the accumulation of traffic on the ground. The taxiing-out traffic can accumulate and reach very high values as aircraft push back, but are not able to take off. In these cases, high values of departing traffic on the ground are due to decreased throughput. Such instances are omitted from the fitting procedure. In Section 2.3.1, we describe a process for filtering out these instances.

The estimation of the data mean regression fit can be formulated as a least-squares problem: Given m pairs of measurements $N(t)$ and $T(t)$, denoted $(u_1, y_1), \ldots, (u_m, y_m)$, we seek a non-

[1] Such cases have been noted during observations at BOS and PHL airports [99, 25].

decreasing, concave function $f_{\text{mean}} : \mathbb{R} \to \mathbb{R}$ that estimates the mean $T = f_{\text{mean}}(N)$. This infinite-dimensional problem is significantly simplified by the fact that N is defined only in the domain of natural numbers ($\mathbb{N}_0$). f_{mean} can be restricted in the domain of $\mathbb{N}_0$ as well, and we need to estimate the values $f_{\text{mean}}(0), f_{\text{mean}}(1), \ldots, f_{\text{mean}}(n)$, where $n = \max(N(t))$. The function f_{mean} is simply a piecewise linear function of N, and the monotonicity and concavity constraints are imposed at the points $0, 1, \ldots, \max(N(t))$ by comparing the values and the slopes of subsequent pieces. f_{mean} is given by the solution to the simple convex optimization problem:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{2.1}$$

subject to:

$$\hat{y}_i = f_{\text{mean}}(u_i), \;\; i = 1, \ldots, m \tag{2.2}$$

$$f_{\text{mean}}(i+1) \geq f_{\text{mean}}(i), \;\; i = 0, \ldots (n-1) \tag{2.3}$$

$$f_{\text{mean}}(i+1) - f_{\text{mean}}(i) \leq f(i) - f(i-1), \;\; i = 1, \ldots (n-1) \tag{2.4}$$

Analogously, the estimation problem of the median throughput as a function of the number of the departure demand is formulated as:

$$\min \sum_{i=1}^{m} |\hat{y}_i - y_i| \tag{2.5}$$

subject to:

$$\hat{y}_i = f_{\text{med}}(u_i), \;\; i = 1, \ldots, m \tag{2.6}$$

$$f_{\text{med}}(i+1) \geq f_{\text{med}}(i), \;\; i = 0, \ldots (n-1) \tag{2.7}$$

$$f_{\text{med}}(i+1) - f_{\text{med}}(i) \leq f_{\text{med}}(i) - f_{\text{med}}(i-1), \;\; i = 1, \ldots (n-1) \tag{2.8}$$

The results of the regression fit can be seen in Figure 2-1b. The mean departure throughput saturates at 11 AC/15 min when $N \geq 22$, while median departure throughput saturates at the same value when $N \geq 19$. We observe that the average departure throughput of this runway configuration under persistent demand is 11 AC/15 min, or 44 AC/hr. Persistent demand is achieved when the number of aircraft taxiing out is around 20.

Finally, we note that this framework can be easily extended to find estimates of upper quantiles

by formulating quantile regression, as described in [72], as a convex optimization problem:

$$\min \sum_{i=1}^{m} \big((1-p)\cdot \max(\hat{y}_i - y_i, 0) + p \cdot \max(-\hat{y}_i + y_i, 0)\big) \tag{2.9}$$

subject to:

$$\hat{y}_i = f_p(u_i), \;\; i = 1, \ldots, m \tag{2.10}$$

$$f_p(i+1) \geq f_p(i), \;\; i = 0, \ldots (n-1) \tag{2.11}$$

$$f_p(i+1) - f_p(i) \leq f_p(i) - f_p(i-1), \;\; i = 1, \ldots (n-1) \tag{2.12}$$

Here, p takes the value of the quantile we are interested in estimating: for the median it is 0.5, for the 90% percentile 0.9, etc.

2.2 Departure throughput as a function of departure demand and arrival throughput

The method described in Section 2.1 can be extended to represent departure throughput as a two variable function of both departure demand and arrival throughput, where the arrival throughput, denoted $A(t)$, is defined to be the number of landings in the 15 minute interval $[t, t+15)$. In other words, we represent the departure throughput, $T(t)$, in the 15 minute interval $[t, t+15)$ as a function of both the departure demand $N(t)$ at time t, and the arrival throughput $A(t)$ in that 15 minute interval. The 2-variable fitting problem has additional constraints that arise from system behavior:

- For a fixed departure demand, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. This is follows from the behavior of capacity envelopes [30], which have been shown to have this property [84].
- For any value of arrival throughput, the departure throughput, as a function of departure demand, cannot increase at a higher rate than for a lower value of arrival throughput.
- For any value of departure demand, the departure throughput, as a function of arrival throughput, cannot decrease at a lower rate than for a lower value of departure demand.

The problem is formulated as follows: Given m triplets of measurements $N(t)$, $A(t)$ and $T(t)$, denoted by $(u_1, v_1, y_1), \ldots, (u_m, v_m, y_m)$, at all times, we seek a function $g : \mathbb{R}^2 \to \mathbb{R}$ that estimates

the mean $T = g(N, A(t))$. Both $N(t)$, $A(t)$ take values in the domain of natural numbers. We only need to estimate the values $g(0,0), g(0,1), \ldots, g(n,l)$, where $n = \max(N(t))$, and $l = \max(A(t))$. Thus, g is a piecewise linear function of $A(t)$ and $T(t)$. The constraints are imposed only between neighboring points, as in the 2D case:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{2.13}$$

subject to:

$$\hat{y}_i = g(u_i, v_i), \;\; i = 1, \ldots, m \tag{2.14}$$

$$g(i+1, j) \geq g(i, j), \;\; i = 0, \ldots (n-1), \forall j \tag{2.15}$$

$$g(i+1, j) - g(i, j) \leq g(i, j) - g(i-1, j), \;\; i = 1, \ldots (n-1), \forall j \tag{2.16}$$

$$g(i, j+1) \leq g(i, j), \;\; j = 0, \ldots (l-1), \forall i \tag{2.17}$$

$$g(i, j+1) - g(i, j) \leq g(i, j) - g(i, j-1), \;\; j = 1, \ldots (l-1), \forall i \tag{2.18}$$

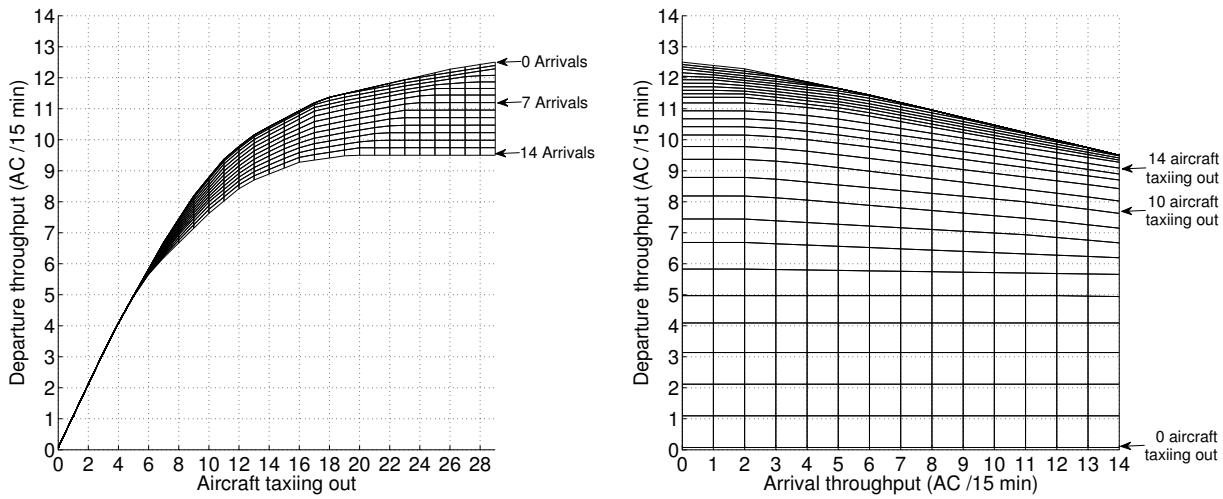
$$g(i+1, j) - g(i, j) \geq g(i+1, j+1) - g(i, j+1), \;\; i = 0, \ldots (n-1), j = 0, \ldots (l-1) \tag{2.19}$$

$$g(i, j) - g(i, j+1) \leq g(i+1, j) - g(i+1, j+1), \;\; i = 0, \ldots (n-1), j = 0, \ldots (l-1) \tag{2.20}$$

Inequalities (2.15) -(2.16) ensure that for a fixed arrival throughput, the departure throughput is a monotonically non-decreasing, concave function of the departure demand. Inequalities (2.17)-(2.18) ensure that for fixed departure demand, the departure throughput is a non-increasing, concave function of the arrival throughput. Finally, Equation (2.19) ensures that the marginal gain in departure throughput from increasing the departure demand increases as the arrival throughput decreases, and Equation (2.20) ensures that the marginal gain in departure throughput from decreasing the arrival throughput decreases as the departure demand decreases.

Under these constraints, the departure throughput is estimated as a function of the departure demand and arrival throughput. Two visualizations of the estimated function g can be seen in Figure 2-2. Figure 2-2a is essentially the mean regression curve of Figure 2-1b parameterized for different levels of arrival throughput. As expected, the arrival throughput impacts the departure throughput. Figure 2-2b displays the same graph from a different perspective: the arrival throughput is the variable on the x-axis and the departure demand is a parameter. We observe that the tradeoff between arrival throughput and departure throughput changes with the departure demand. We also note that for high values of departure demand, the departure runway(s) are under persistent

demand and so the curves for values of N coincide and envelop all average departure throughput data points. Thus, this envelope can be also interpreted as the capacity envelope for this runway configuration: It shows the average departure throughput under high departure demand as a function of arrivals.



(a) Departure throughput as function of departure demand.

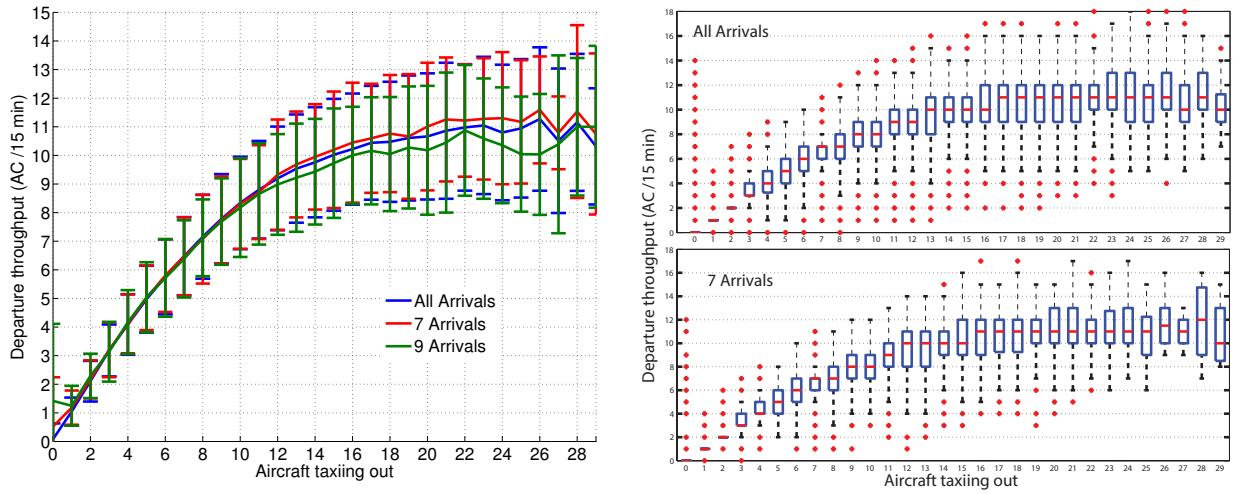
(b) Departure throughput as function of arrival throughput.

Figure 2-2: BOS departure throughput in configuration (VMC; 22L, 27 | 22L, 22R) as function of arrivals and departure demand.

2.2.1 Variance of departure throughput

It could be hypothesized that the high variance of the departure throughput, which can be observed in Figure 2-1a, is explained by the (hidden) arrival throughput variable, and that by controlling for the arrival throughput variable, the variance of the departure throughput would decrease significantly. In order to informally inspect this hypothesis, we plot the departure throughput as a function of departure demand for two frequently observed values of arrival throughput, overlaid with the departure throughput for all values of arrival throughput in Figure 2-3a. We observe that the variance of the departure throughput remains high even at individual arrival throughput values. In Figure 2-3b, we visualize the boxplots of the throughput measurements grouped for every value of the departure demand for only 7 arrivals, and for all arrivals. From the boxplots, we observe that the 25th and 75th percentiles are spread over a range of 3 AC/15 min for both cases under high departure demand.

These results suggest that the observed variance of the departure throughput must be explained



(a) Mean and standard deviation of departure throughput as function of departure demand for different numbers of arrivals.

(b) Boxplots of the departure throughput as function of departures demand for different numbers of arrivals.

Figure 2-3: BOS departure throughput in configuration (VMC; 22L, 27 | 22L, 22R) for different numbers of arrivals.

by other variables, such as, fleet mix, downstream restrictions, human factors (controller performance, pilot response times), and unexpected incidents (runway closures, mechanical failures, etc.).

2.3 Operational throughput envelope

2.3.1 Data filtering

Traditional statistical methods for predicting the response variable (departure throughput) as a function of several independent variables (departure demand, arrival throughput, fleet mix) do not exploit the structure of the problem and do not impose the constraints that result from the physics of the system. For this reason, we follow a different approach: We first isolate instances of high departure demand and runway availability, and then estimate the departure throughput as a function of the arrival throughput. The combination of the independent variable (arrival throughput) and the dependent variable (departure throughput) yield the airport capacity. For distinguishing it from other measures of capacity, we call this quantity "operational throughput" .

To this end, we need to estimate the threshold N^* at which the departure throughput stops varying with departure demand. There are several ways to estimate N^*, for example, through the inspection of Figure 2-1a or Figure 2-1b. Both plots suggest that the mean throughput saturates at $N^* = 22$.

A more robust way to identify N^* is the following: The departure throughput is a function of the departure demand and the arrival throughput. We construct a regression tree that represents the dependence of the response variable (departure throughput) on the two exogenous independent variables (departure demand and arrival throughput). A simplified version of the resulting regression tree for this example is shown in Figure 2-4.

The regression tree enables us to represent the nonlinear relation between the departure throughput and the two independent variables. For example, if the arrival throughput made use of all the capacity of the airport, the departure throughput would be zero independent of the value of the departure demand. At the other extreme, if the arrival demand is zero, all the capacity is allocated to departures, and a high value of demand might be required to sustain the departure throughput at its maximum value. From the regression tree, we identify the node at which the highest value of departure demand explains some of the variation of the departure throughput. For this example, this node is $N \geq 22$, as can be seen as can also be seen in Figure 2-4. This means that for departure demand greater than or equal to 22, the departure demand does not explain the variation of the departure throughput independent of the value of the arrivals. Thus, $N^* = 22$.

To avoid over-fitting, the regression trees are pruned with 10-fold cross-validation. Although this method has certain shortcomings, such as the assumption of normally distributed errors in each node and the instability of the pruned trees, it offers a simple and intuitive representation of the different demand, arrival throughput and departure throughput scenarios. We have found that it works well in different runway configurations as well as airports.

In addition, it is possible that some very high values of N are due to an operational anomaly, as explained in Section 2.1. These instances can be identified from the regression tree: At some nodes, an increasing value of N would lead to a decreasing value of departure throughput. In the present example, $N \geq 27$, $A < 9$ represents such a node, as can be seen in Figure 2-4. These instances are filtered out and excluded from further analysis. We denote by $N_{\max}$ the highest value of N for which the throughput does not change with N. Here, $N_{\max}$=26.

Next, we use the data points for which $N^* \leq N \leq N_{\max}$ to isolate the instances of persistent departure demand. However, as explained earlier, persistent departure demand does not necessarily guarantee runway availability. For example, there could be a 15-minute interval during a busy high-demand period when the runway is temporarily unavailable due to a thunderstorm, or a safety incident. Unfortunately, this information is not recorded in the ASPM data, and does not reflect runway capacity. One way to reduce this bias is to exclude the 15-minute intervals during which the

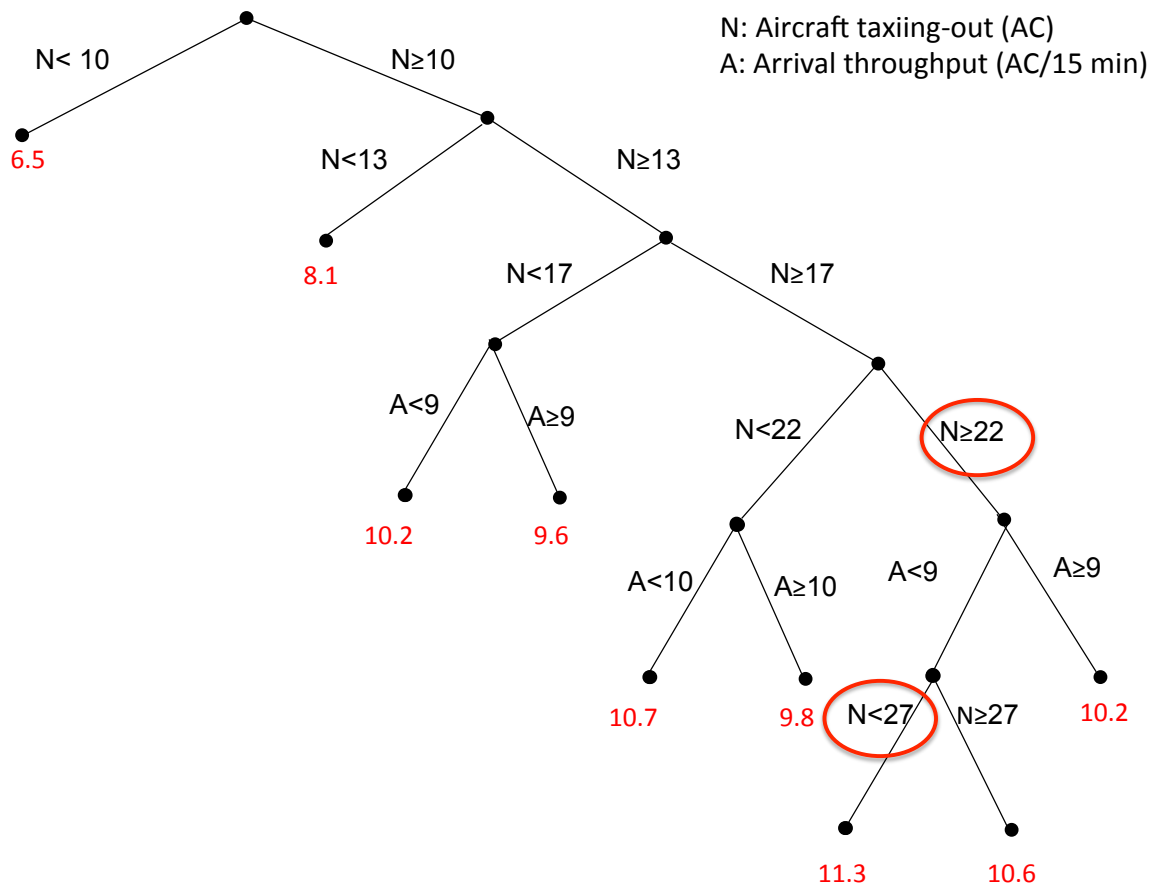


Figure 2-4: BOS departure throughput (AC/15 min) as a function of arrival throughput (A) and departure demand (N).

total departure throughput is extremely low. The threshold used to determine acceptable values of departure throughput is set to be the highest Heavy aircraft departure throughput recorded in any 15-minute interval. This threshold is reasonable because Heavy aircraft require the longest inter-departure separation times. For the case of the 22L, 27 | 22L, 22R configuration at BOS, it is 5 Heavy aircraft departures. Our assumption is that if there was a time interval during which 5 Heavy aircraft departures were serviced (in addition to other departures and arrivals), 5 departures in total should be feasible at any time interval, regardless of fleet mix, or arrival throughput. We have found that this simple filtering works sufficiently well in all airport environments considered. In Chapter 3, we discuss how the filtering can be improved in the presence of information on weather and airspace constraints.

An important concept in both theoretical and empirical capacity envelopes determination is that of the *arrival priority* capacity, which is the statistically significant, highest value of arrival

throughput that can be realized at the airport. In order to estimate this quantity, we follow the procedure outlined by Ramanujam and Balakrishnan [97]. All data points with arrival throughput values higher than the estimated *arrival priority* capacity are filtered out as outliers.

In the remaining discussion of the fitting and estimation problems, we refer to the dataset of the airport in saturation after having removed extreme values of N, A, and T, as the *filtered dataset in saturation*.

2.3.2 Estimation method

Given k pairs of measurements $A(t)$ and $T(t)$, denoted $(v_1, y_1), \ldots, (v_k, y_k)$ from the filtered dataset in saturation, we seek a non-increasing, concave function $h : \mathbb{R} \to \mathbb{R}$ that estimates the mean $T(t) = h(A(t))$. Similarly to the previous cases, we only need to estimate the values $h(0), h(1), \ldots, h(l)$, where $l = \max(A(t))$. Thus, function h is a piecewise linear function of A and the monotonicity and concavity constraints are imposed at the points $0, 1, \ldots, l$ by comparing the values and the slopes of subsequent pieces. The formulation of this estimation problem is as follows:

$$\min \sum_{i=1}^{k} (\hat{y}_i - y_i)^2 \tag{2.21}$$

subject to:

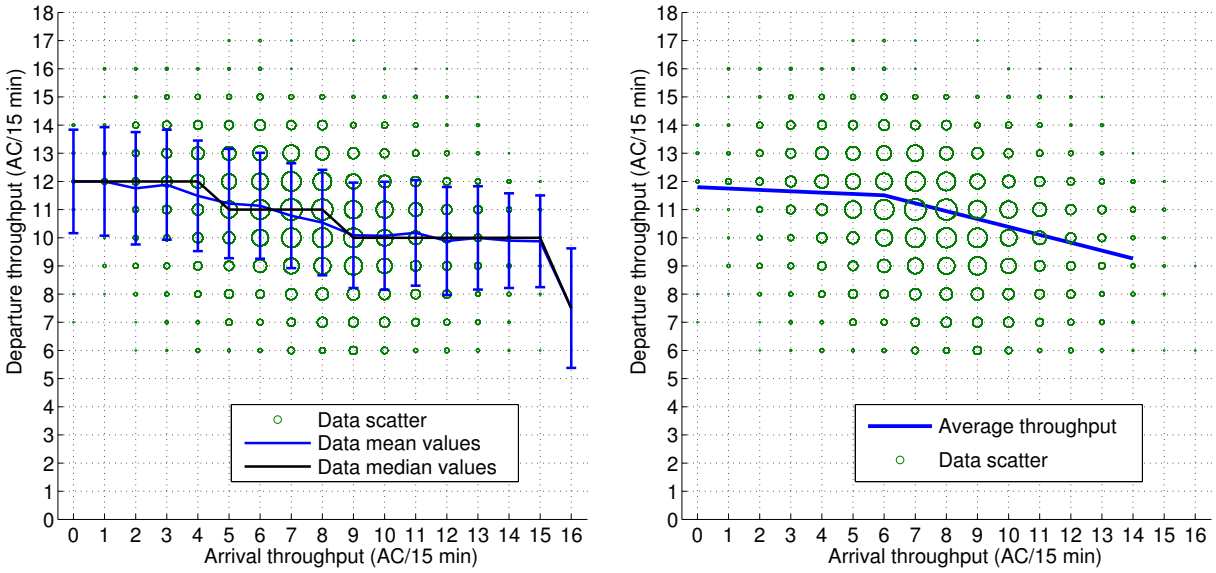
$$\hat{y}_i = h(v_i), \;\; i = 1, \ldots, k \tag{2.22}$$

$$h(i+1) \leq h(i), \;\; i = 0, \ldots (l-1) \tag{2.23}$$

$$h(i+1) - h(i) \leq h(i) - h(i-1), \;\; i = 1, \ldots (l-1) \tag{2.24}$$

The mean and median of the departure throughput, for each value of the arrival throughput, can be seen in Figure 2-5a, and the corresponding fitted function is shown in Figure 2-5b. The plot of Figure 2-5b provides an estimate of the average departure throughput as a function of the arrival throughput under persistent departure demand and runway availability. We refer to this plot as the *operational throughput envelope*.

Comparing Figures 2-5a and 2-5b, we note that the values in the raw measurements dataset of Figure 2-5a include as many as 16 AC/15 min. However, the *arrival priority* capacity is estimated to be 14 AC/15 min for this runway configuration and for this reason the operational throughput envelope is estimated for values of arrivals between 0 and 14 AC/15 min. The data points with more than 14 AC/ 15 min are filtered out as outliers and are not considered in the estimation process.



(a) Mean, median, and standard deviation of the departure throughput for all values of the arrival throughput.

(b) Regression of the departure throughput as function of the arrival throughput.

Figure 2-5: BOS operational throughput envelope in configuration (VMC; 22L, 27 | 22L, 22R).

In agreement with the theoretical models [30], the *arrival priority* data points can accommodate additional departures. In this case, the total operational throughput at this operating point is 14 arrivals and 9 departures. This means, that under arrival priority capacity, up to 9 departures/15 min can be accommodated.

2.4 Parametrization of the operational throughput envelope

2.4.1 Role of ATC in allocating capacity

In this section, the operational throughput envelope is further parametrized in order to investigate other significant capacity tradeoffs, besides the tradeoff between departures and arrivals. In all case studies we focus on departure capacity for the following reasons:

Firstly, arrival capacity is an exogenous variable from the perspective of the airport. Arriving traffic is handled by the Terminal Radar Approach Control Facility (TRACON) and is routed to the airport according to allocated resources or agreed rates. For example, in the case of the 22L, 27 | 22L, 22R configuration at BOS, the highest number of landings, as sequenced by the

TRACON, is 15 AC/ 15 min [2] independent of the departure rate. One can therefore analyze the arrival throughput independent of the departure throughput, and study its dependence on runway configuration, arrival demand, fleet mix, weather etc. One such example would be a plot showing how the arrival capacity changes with route availability, or with the number of Heavy aircraft in the arrival mix.

Secondly, from the perspective of the Airport Traffic Control Tower Facility (ATCT), what matters is the departure capacity, since it have no direct control over the arrival capacity. Thus, for the ATCT, it is important to measure how arrivals impact departure operations. The fact that the arrival throughput is fixed by an external facility (the TRACON, an Area Control Center, or FAA's Command Center), and departures are realized as external conditions (including arrivals) permit, justifies the use of the arrival capacity as the independent variable in the capacity envelope representation.

An alternate method would simultaneously predict both arrival and departure capacity given some interdependent variables: Arrival fleet mix, spacing and sequence, departure fleet mix, route availability, winds etc. Such an analysis could be performed with multivariate multiple regression or multivariate regression trees, and would estimate the average capacity given a set of exogenous conditions. However, this analysis would not yield the simple representation of Figure 2-5b.

2.4.2 Multi-variable statistical models of departure capacity

In this section, we continue working with the filtered dataset in saturation. We address the problem of estimating the departure throughput as a function of the arrival throughput and the fleet mix. The fleet mix is not a simple numerical variable like the number of arrivals, and its impact is highly dependent on the design and operating procedures of each airport (such as runway configuration, sequencing decisions, airspace design, and local procedures). For runway configuration 22L, 27 | 22L, 22R under VMC at BOS, given operational information from controllers and observations in the ATC tower, our hypothesis is that the fleet mix can be represented with two variables:

- Number of propeller-powered aircraft, or props, in the mix that is taking off in the 15-minute interval (denoted P_{Deps}) .
- Number of Heavy aircraft, or Heavies in the mix that is taking off in the 15-minute interval (denoted H_{Deps}) .

[2] Interviews with Controllers from Boston TRACON and Boston ATCT

Props are given dispersal headings when sequenced between jet departures, and are thus expected to increase the departure throughput. Heavy aircraft introduce longer separation requirements, and are expected to decrease the departure throughput. Based on these facts, we model the response variable, departure throughput ($T(t)$) in each 15-minute time interval $[t, t+15)$, as a function of four potential explanatory variables:

1. Number of aircraft ($N(t)$) taxiing out at time t.[3]
2. Arrival throughput ($A(t)$) in the 15-minute interval.
3. Number of props ($P_{Deps}(t)$) departing in the 15-minute time interval.
4. Number of Heavy aircraft ($H_{Deps}(t)$) departing in the 15-minute time interval.

We analyze the relation between all the variables in the model using Mutual Information (MI) and locally weighted scatterplot smoothing (lowess) regression. Mutual Information can be applied to study the relation between two random variables, X and Y. It measures how much of the uncertainty of X is reduced when Y is observed, but does not capture situations in which two random variables Y and Z combined reduce the uncertainty of X. For discrete random variables X and Y, MI is expressed as:

$$I[X;Y] = \sum_x \sum_y p_{xy}(x,y) \log \frac{p_{xy}(x,y)}{p_x(x)p_y(y)} \tag{2.25}$$

All the random variables considered here are discrete and have a small domain (between 0 and 30), and their probability distributions are approximated by empirical ones. The MI scores between each potential explanatory random variable considered and the departure throughput is shown in Table 2.1. We also calculate and report the correlation coefficient.

From Table 2.1 we note that the random variables P_{Deps} and A have the highest mutual information and linear correlation with departure throughput. As expected, departure demand has very low mutual information with departure throughput. Finally, the Heavy departures appear to not adversely impact departure throughput.

For lowess regression, we use the *pairs* function of the R programing language [96], which produces panels with correlations between all variables of the model. Each panel shows the scatterplot

[3]Although we have established that the departure throughput does not change significantly with the number of aircraft on the ground in the filtered dataset in saturation, it is useful to revisit this hypothesis in the multi-variable model.

Table 2.1: Mutual information and correlation between departure throughput and potential explanatory variables.

Variable	Mutual Information with departure throughput (T)	Correlation with departure throughput (T)
Departure demand (N)	0.023	0.002
Arrival throughput (A)	0.079	-0.245
Departing props (P_{Deps})	0.151	0.414
Departing Heavies (H_{Deps})	0.019	0.045

between the variable on the vertical axis and the variable on the horizontal axis as well as the lowess curve, in red color, through the set of data points. Lowess fits follow the general trend of the data and are a good measure of the correlation between the two variables [27]. The response variable, departure throughput, is shown on the y-axis of the top row of the panels.

From Figure 2-6, we observe that:

- The lowess fit line for the variable pair (N, T) does not exhibit any large or systematic deviation from the $T = 11$ line. This is further evidence that N^* was calculated correctly, and that for $N \geq N^*$, there is no correlation between the departure demand and the departure throughput. The departure throughput is shown to be stable at 11 AC/15 min, the same value that was calculated using the estimation method of Figure 2-1b.

- The lowess fit line for the variable pair (A, T) follows the same trend as in Figure 2-5b: It shows that the departure throughput decreases from 12 to 9 AC/15 min as a concave function of the arrival throughput.

- The lowess fit line for the variable pair (H_{Deps}, T) exhibits a rather inconclusive trend. The curve initially increases from 10.5 to 11 and then decreases and stabilizes at around 10 AC/15 min.

- The lowess fit line for the variable pair (P_{Deps}, T) exhibit a clear positive correlation between the two: As the number of prop departures increases from 0 to 6, the total departure throughput appears to increase from 9 to 14 AC/15 min.

The relationship between the departure throughput and these four variables is also examined with other statistical tools, such as, regression trees, random forests and generalized additive models. They all lead to the same conclusion: The two most significant explanatory variables are the

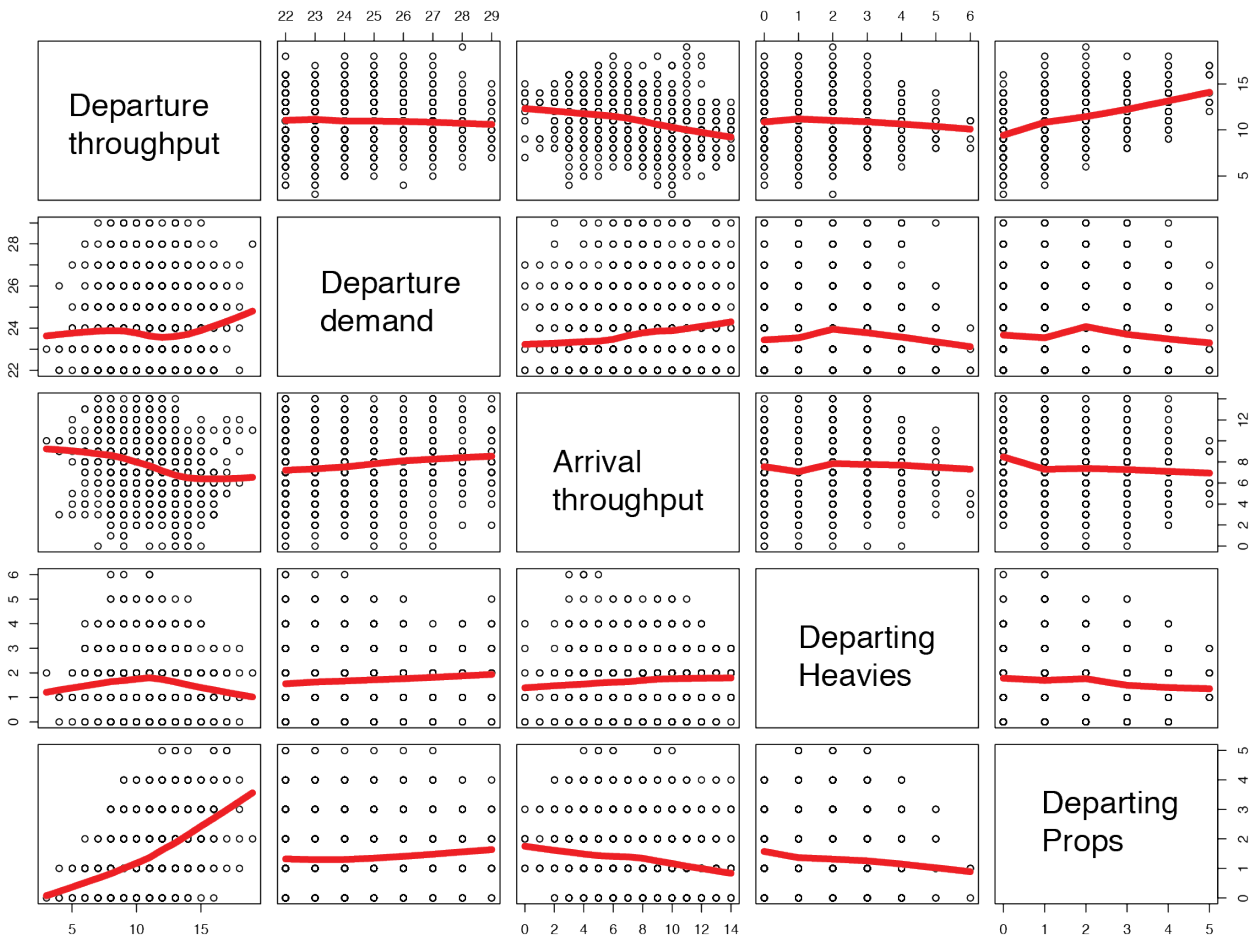


Figure 2-6: Correlation between departure throughput, departure demand, arrival throughput, Heavy departures and departing props.

arrival throughput and the prop departures. Heavy aircraft departures do not impact the departure throughput significantly. While surprising, this result can be explained by the operational procedures at BOS. Controllers use the large wake vortex separation requirement between a Heavy aircraft and a subsequent departure to perform runway crossings, diminishing the impact of Heavy departures on throughput.

2.4.3 Operational throughput parametrized by fleet mix

Having established that in the filtered dataset in saturation, the departure throughput is primarily a function of arrival throughput and prop departures, we estimate the departure throughput as a function of arrival throughput and prop departures using an approach similar to the one described in Section 2.2. Similarly to the other curve fitting problems, we exclude data points with extreme values of the variable P_{Deps} by filtering out datapoints, which exceed the top 1 percentile of the

measured values of P_{Deps}.

Given k triplets of measurements $A(t)$, $P_{Deps}(t)$ and $T(t)$, denoted by $(v_1, w_1, y_1), \ldots, (v_k, w_k, y_k)$, in the filtered dataset in saturation, we seek a function $h_p : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $T(t) = h_p(A(t), P_{Deps}(t))$. Again, we only need to estimate the points $h_p(0,0), h_p(0,1), \ldots, h_p(l,m)$, where $l = \max(A(t))$ and $m = \max(P_{Deps}(t))$. Thus, the function h_p is a piecewise linear function of $A(t)$ and $P_{Deps}(t)$. The constraints are imposed only between neighboring points, as was done in the 2D case:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{2.26}$$

subject to:

$$\hat{y}_i = h_p(v_i, w_i), \;\; i = 1, \ldots, k \tag{2.27}$$

$$h_p(i+1, j) \leq h_p(i, j), \;\; i = 0, \ldots (l-1), \forall j \tag{2.28}$$

$$h_p(i+1, j) - h_p(i, j) \leq h_p(i, j) - h_p(i-1, j), \;\; i = 1, \ldots (l-1), \forall j \tag{2.29}$$

$$h_p(i, j+1) \geq h_p(i, j), \;\; j = 0, \ldots (m-1), \forall i \tag{2.30}$$

$$h_p(i, j+1) - h_p(i, j) \leq h_p(i, j) - h_p(i, j-1), \;\; j = 1, \ldots (m-1), \forall i \tag{2.31}$$

$$h_p(i+1, j+1) - h_p(i+1, j) \leq h_p(i, j+1) - h_p(i, j), \;\; i = 0, \ldots (l-1), j = 0, \ldots (m-1) \tag{2.32}$$

$$h_p(i, j+1) - h_p(i+1, j+1) \geq h_p(i, j) - h_p(i+1, j), \;\; i = 0, \ldots (l-1), j = 0, \ldots (m-1) \tag{2.33}$$

Inequalities (2.28)-(2.29) are analogous to those in the case of the capacity envelope (Inequalities (2.23)-(2.24)). For a given fleet mix, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. Inequalities (2.30)-(2.31) ensure that for fixed arrival throughput, the departure throughput is a non-increasing, concave function of the number of prop departures. This constraint models the operational observation that increasing the number of props is expected to boost departure throughput. It is also expected to deliver diminishing gains as the number increases, because opportunities for dispersal headings decrease.

Similarly, Equation (2.32) ensures that the marginal gain in departure throughput from increasing the number of props by one unit decreases as the arrival throughput increases. The operational reason for this is that as the number of arrivals increases, there is more pressure to cross arriving aircraft on runway 22R, and this pressure can reduce the impact of dispersal headings. The runway is likely to be utilized for crossing arriving aircraft during inter-departure intervals independent of

the separation requirements.

Finally, Equation (2.33) ensures that the marginal gain in departure throughput from decreasing the arrival throughput by one unit increases as the number of prop departures increases. If decreasing the arrival throughput by one unit enables the airport to increase departure throughput by some amount, decreasing the arrival throughput by one unit and replacing one jet aircraft with one prop will lead to at least the same improvement in departure throughput.

The plot of the estimated function, $h_p(A, P_{Deps})$, can be seen in Figure 2-7 overlaid with the dashed curve of Figure 2-5b (*average throughput*). The comparison shows that the solid lines in Figure 2-7 are, in fact, the dashed line parameterized by the number of props departing in that 15-minute interval.

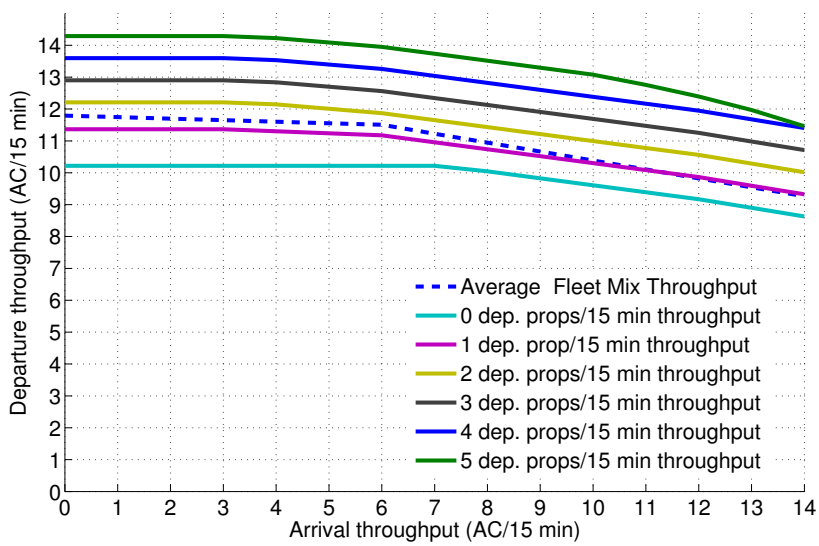


Figure 2-7: BOS parametrized operational throughput envelope in configuration (VMC; 22L, 27 | 22L, 22R).

We observe the following features in Figure 2-7:

- The average departure throughput curve lies between those corresponding to a fleet mix of 1 departing prop/15 min and 2 departing props/15 min, which is consistent with the number of props in the fleet mix at BOS (around 15% in 2007).

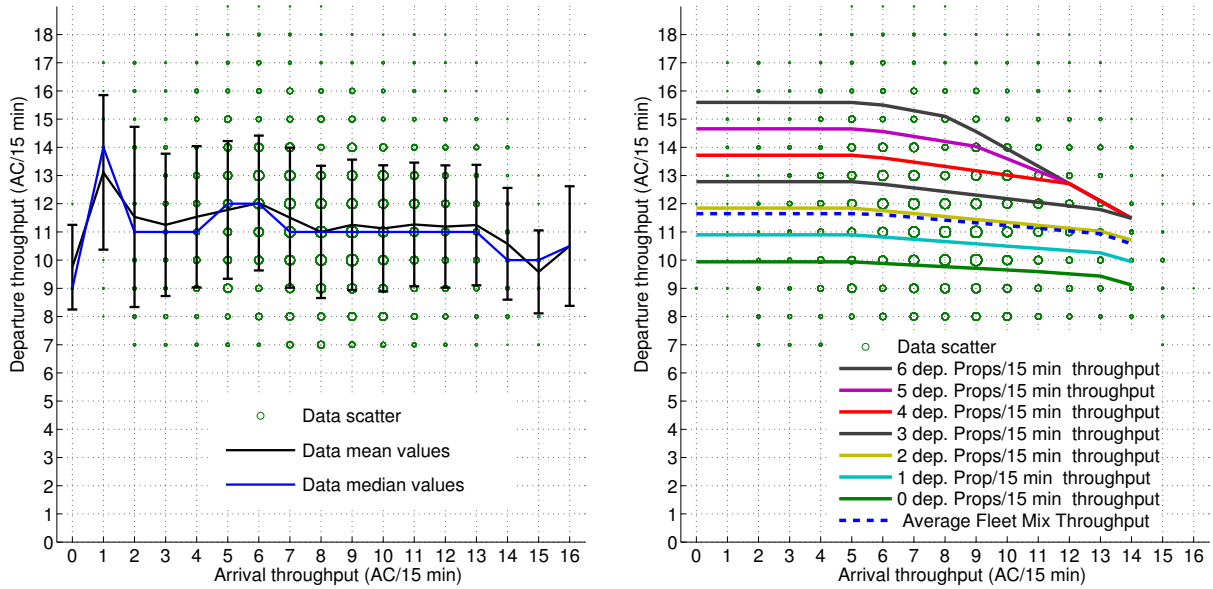
- The number of props has a significant impact on the departure throughput. During the most common operating scenarios in which the arrival throughput is 5-10 aircraft/15 min and the number of prop departures is 0-2/15 min, the departure throughput increases at a a rate of almost one aircraft for each additional prop.

- From this plot and the previous statistical analyses, we conclude that for this runway configuration at BOS, the fleet mix is a more significant explanatory variable for the departure throughput than the arrival throughput is. The departure throughput decreases with the arrival throughput by at most 2.6 AC/15 min, for an increase of arrival throughput from 0 to 14 AC/15 min. In contrast, increasing the number of props in the fleet mix from 0 to 5 increases the departure throughput by 4.4 AC/15 min.

For completeness, we provide the parametrized operational throughput envelopes for the other major runway configurations at BOS under VMC, 4R, 4L | 4R, 4L, 9 and 27, 32 | 33L, in Figures 2-8 and 2-9. We observe that the three runway configurations have similar characteristics: In all of them arriving traffic utilizes two arrival runways and has the same *arrival priority* capacity value, 14 AC/15 min. This is very close to the FAA airport arrival rates (AAR)[4], which are 61 arrivals/hr and 59 arrivals/hr for the configurations 4R, 4L | 4R, 4L, 9 and 22L, 27 | 22L, 22R [39] correspondingly. There is a fraction of unutilized arrival capacity, since the empirical capacity is estimated at 14 AC/15 min (or 56 AC/hr), which is 3-5 aircraft fewer than the AAR . For configuration 27, 32 | 33L at VMC, the AAR is 44 arrivals/ hour, which is much smaller than the estimated arrival capacity (56 AC/hr). However, Runway 32 is exempt from Traffic Management Initiatives (TMI's) in the published AAR for this configuration [39]. Its empirical arrival priority capacity can therefore be much higher, and is 56 AC/hr according to our analysis.

We also note that props increase the departure throughput in a similar fashion in all runway configurations: As the number of props increases from 0 to 3, the departure throughput increases by 2 AC/15 min. A policy implication of this observation is that the airport should incentivize the use of props as opposed to jets of similar size, as this increases overall passenger capacity. By the same rationale, the marginal external cost of a prop departure is much smaller than the marginal cost of a jet departure.

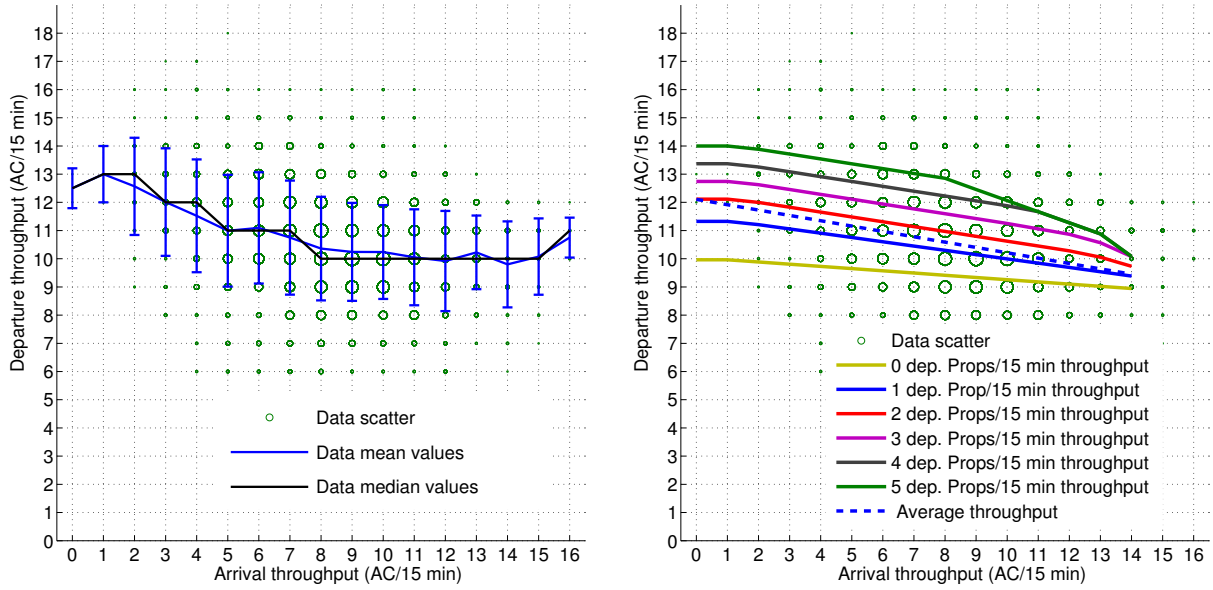
[4]the number of arrivals an airport is capable of accepting each hour



(a) Mean, median, and standard deviation of the departure throughput for all values of the arrival throughput.

(b) Departure throughput as a function of arrival throughput and departing props.

Figure 2-8: BOS operational throughput envelope in configuration (VMC; 4R, 4L | 4R, 4L, 9).



(a) Mean, median, and standard deviation of the departure throughput for all values of the arrival throughput.

(b) Departure throughput as a function of arrival throughput and departing props.

Figure 2-9: BOS operational throughput envelope in configuration (VMC; 27, 32 | 33L).

2.4.4 Comparison to empirical capacity envelopes

In the previous section, we showed that the departure throughput can be modeled as a function of the arrival throughput and the fleet mix under persistent departure demand. For the case of BOS configuration (22L, 27 | 22L, 22R), we saw that the average departure throughput under persistent demand (11 AC/15 min) takes values in the range of 8.5 to 14.5 AC/15 min depending on the arrival throughput and the fleet mix (Figure 2-7).

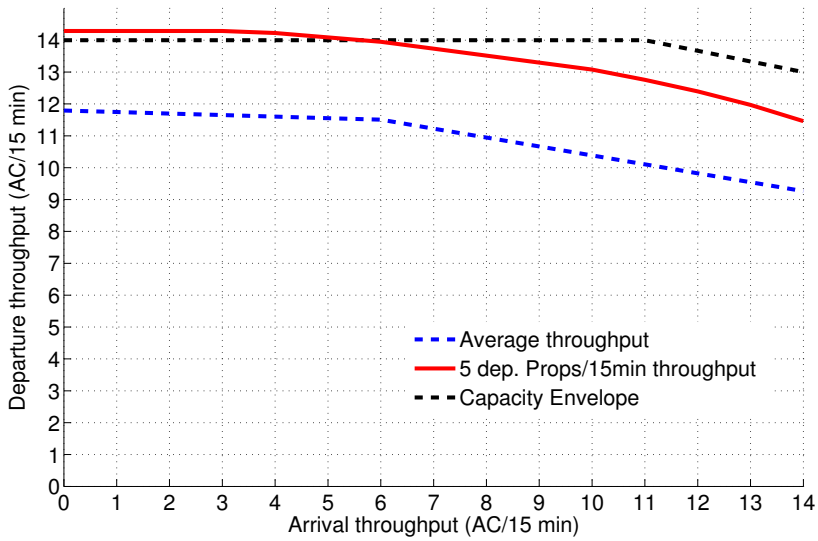


Figure 2-10: Operational throughput envelope and capacity envelope comparison for BOS configuration (VMC; 22L, 27 | 22L, 22R).

In this section, we compare the estimated functions plotted in Figure 2-7 to empirical capacity envelope estimates. Empirical capacity envelopes represent the highest departure throughput as a concave non-increasing function of the arrival throughput: The curve enveloping the observed maximum arrival and departure counts, after correcting for outliers, is considered the airport capacity envelope [49]. For the capacity envelope estimation, we use the approach proposed by Ramanujam and Balakrishnan [97], which for this runway configuration at BOS yields the capacity envelope plotted in Figure 2-10. The maximum total capacity is achieved at the point of *arrival priority* capacity: 27 movements/15 minutes (14 arrivals and 13 departures).

The capacity envelope is overlaid with the average departure throughput curve and the highest departure throughput curve (the one with the most favorable fleet mix). We observe that the capacity envelope is close to the throughput curve of the most favorable fleet mix. This reveals an inherent ambiguity in the capacity envelope analysis. While the commonly used definition of

capacity is "the *average* number of movements that can be performed on the runway system in the presence of continuous demand" [30], most empirical capacity envelope estimation methods focus on the best-case scenario, that is, the *maximum* number of movements that can be performed on the runway system in the presence of continuous demand [49]. These maximum counts are achievable only under special circumstances, such as a favorable fleet mix or a favorable sequence. In the case of BOS, the difference between the *capacity envelope* and the *operational throughput envelope* estimates for the departure capacity at a given value of arrival throughput can be as large as 4 AC/15 min, as can be seen in Figure 2-10.

2.5 Predictive ability of proposed methods

The proposed estimation procedure can be viewed as a supervised learning method, since we can select independent variables and use them to predict the dependent variable. In this process, we apply constraints that result from the physics of the system to avoid over-fitting. However, it is still useful to measure the predictive power of the proposed method. We propose to use five measures for goodness-of-fit: The mean absolute error (MAE), the mean square error (MSE), the root mean square error (RMSE), the mean absolute percentage error (MAPE) and the root mean square percentage error (RMSPE).

$$MAE = \frac{1}{N}\sum_{i=1}^{m}|\hat{y}_i - y_i| \tag{2.34}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \tag{2.35}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2} \tag{2.36}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{m}|\frac{\hat{y}_i - y_i}{y_i}| \tag{2.37}$$

$$RMSPE = \sqrt{\frac{1}{N}\sum_{i=1}^{m}(\frac{\hat{y}_i - y_i}{y_i})^2} \tag{2.38}$$

The estimation procedure minimizes the MSE given the constraints. However, this measure penalizes higher errors disproportionally compared to small errors. For this reason, we also record the MAE, which linearly penalizes all errors. The RMSE offers a more direct comparison to the

MAE. In order to assess our results, we compare three different estimates for $\hat{y}_i$:

1. The arithmetic mean (AM) of the departure throughput: $y = \frac{1}{m}\sum_{i=1}^{m} y_i$.

2. The predicted throughput predicted from the operational throughput envelope (OTE): $h(v_i)$.

3. The predicted throughput from the parametrized throughput envelope (PTE): $h_p(v_i, w_i)$.

The MSE, when the estimate is the (AM) of the departure throughput is an estimate of the variance of the departure throughput. Comparing it to the MSE from the estimates of the (OTE) and the (PTE) indicates how much of the variation of the departure throughput is explained by the arrival throughput and the arrival throughput and prop departures, respectively.

Table 2.2: Statistics of the training and test datasets.

Dataset	Time frame	Source	Size	Use	Mean	Variance
1	Jan-Dec 2007	ASPM	9623	Train	10.56	3.72
2	Jan-Dec 2008	ASPM	7429	Test	10.56	3.56
3	Jan-Dec 2011	ASDE-X	5970	Train	11.45	2.79
4	Nov-Dec 2010	ASDE-X	628	Test	11.35	2.23

All the datasets described in Table 2.2 consist of the data points in saturation, that is the data points at which $N \geq N^*$. When training the (OTE) and (PTE) estimators, appropriate filtering is applied as described in Section 2.3. However, the resulting predictors are applied to the whole dataset in saturation in both the training and test datasets, by approximating the extreme values of independent variables with the closest ones in the filtered dataset. Dataset 1, the first training dataset, is the one we have been working with so far, namely ASPM data from 2007 in saturation. Dataset 2, the dataset for which we test the predictors estimated with Dataset 1, is the corresponding one for 2008. The saturation point is the same one as in the training dataset, that is $N^* = 22$. Comparing the ASPM datasets from 2007 and 2008 in Table 2.2, we notice that they have the same average departure throughput: 10.56 AC/15 min. The variance of the departure throughput is higher in 2007. In addition, for 2007, we have more data points, since it was a year with more aircraft movements and more congestion at BOS. As far as the predictions are concerned, we note from Table 2.3 that both the (OTE) and the (PTE) reduce the unexplained variation in the dataset for both years. However, the benefit offered by the (OTE) is rather insignificant, since the mean absolute error is reduced from 1.53 to 1.48 in the test dataset. This was also noted when parametrizing the saturation plots for the arrival throughput (Figure 2-3). However, in both

the training and test data, the (PTE) offers a significant reduction in both the MAE and MSE, confirming that fleet mix is a very significant factor when studying the capacity of BOS.

Finally, for completeness, we repeat the same procedure with ASDE-X data, a dataset obtained from ground surveillance data [37]. ASDE-X is a less noisy, more detailed and more complete dataset. It is obtained from ground surveillance, does not rely on aircraft reporting of events, and includes more flights. Demand is measured more accurately. We only include aircraft on the active movement area when measuring demand, and we exclude aircraft in holding areas. Similarly, we can recognize events like aborted take-offs. Finally, we can infer exact takeoff times, with seconds precision[5]. By contrast, ASPM times are approximated from ACARS messages [88]. For these reasons, we note with respect to Table 2.2, that the average throughput has both a higher value and a smaller variance. The (PTE) for ASDE-X, shown in Figure 2-11, has a higher predictive power in both the training and the test dataset. As can be seen from Table 2.3, the MSE is reduced from 2.79 to 1.92 and from 2.24 to 1.20, in the training and the test datasets, respectively.

Despite its significant benefits, ASDE-X is a very new data source, and is available only for a few airports, and only recently. It also requires significant filtering effort in order to obtain useful flight event information. This difficulty can be observed in Table 2.2, where we only have two months of ASDE-X data for testing our estimators. Runway 33 was closed for most of 2011, thus we cannot construct the OTE and PTE for 27, 32 | 33L using the training dataset. These difficulties lead us to focus our analysis on ASPM and show that significant information can be extracted despite its limitations. For example, Figures 2-11a and 2-11b offer qualitatively the same information as Figures 2-5b and 2-8b. The balanced operations capacity of runway configuration 22L, 27 | 22L, 22R when measured with ASPM is (10,10), whereas it is (11,11) when measured with ASDE-X. Similarly it is (11,11) for 4R, 4L | 4R, 4L, 9 with ASPM and (12, 12) with ASDE-X. Regardless of the data source, we can conclude that (4R, 4L | 4R, 4L, 9) is 20% more efficient than 22L, 27 | 22L, 22R. Similarly, for runway configuration 22L, 27 | 22L, 22R, both data sources offer the same insights concerning the role of the props, that is increasing their number from 0 to 3 in the departing fleet mix increases the departure throughput by 2 AC/15 min under the most frequently encountered arrival rates (5-12 AC/15 min).

[5]The filtering and pre-processing of ASDE-X data was conducted by H. Khadilkar with the methods explained here [68]

Table 2.3: Comparison of three different estimators for the departure throughput on the training and the test datasets.

Dataset	Estimator	MAE	MSE	RMSE	MAPE	RMSPE
1	AM	1.56	3.72	1.93	16.5%	22.3%
1	OTE	1.51	3.56	1.89	15.8%	21.3%
1	PTE	1.36	2.92	1.71	14.0%	19.3%
2	AM	1.53	3.56	1.89	15.7%	21.4%
2	OTE	1.48	3.49	1.87	15.3%	21.3%
2	PTE	1.35	2.93	1.71	14.01%	20.0%
3	AM	1.32	2.79	1.67	13.0%	16.0%
3	OTE	1.29	2.72	1.65	11.7%	15.8%
3	PTE	1.09	1.92	1.39	9.9%	13.4%
4	AM	1.19	2.24	1.50	10.7%	13.6%
4	OTE	1.16	2.16	1.47	10.5%	13.5%
4	PTE	0.87	1.20	1.10	7.7%	9.95%

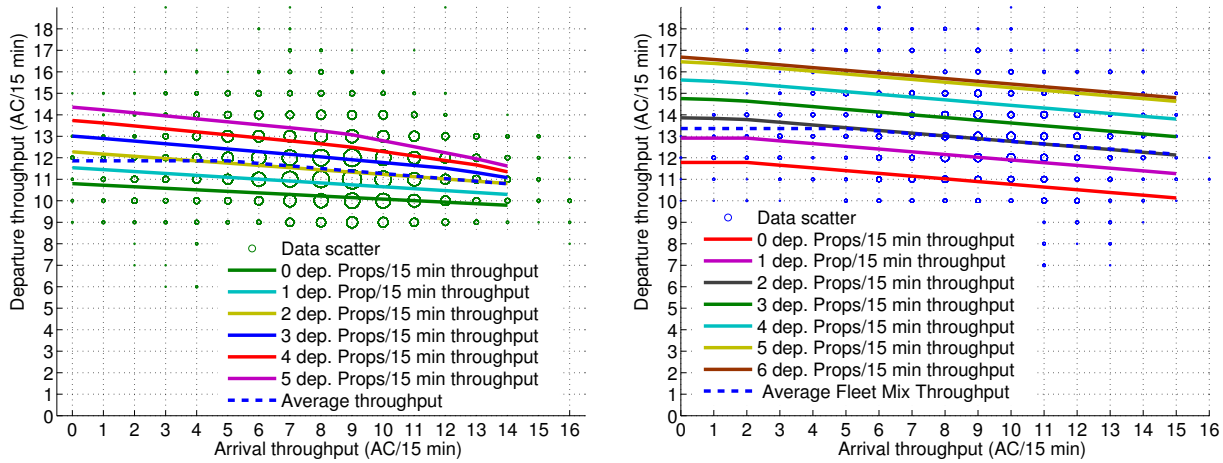
2.6 Jet departure throughput at BOS

In Section 2.4, we showed that the departure throughput in the major VMC configurations at BOS increases at the rate of almost one operation per prop in the departure fleet mix. This observation prompts us to investigate the hypothesis (of the BOS ATCT) that jet and prop departure operations are fairly decoupled at BOS. For studying the interaction of jets and props in more detail, we further adapt the proposed methodology. We also show how the proposed framework can be used to study second-order effects not revealed with standard statistical analysis tools. In particular, we examine the impact of Heavy aircraft departures in more detail, although according to the statistical analysis in Section 2.4, Heavy aircraft do not significantly impact the departure throughput of BOS.

2.6.1 Interactions between jet and prop departures

First, we estimate jet throughput under persistent jet demand as a function of arrival throughput and prop departures. For this, we use the filtered dataset in saturation, but for jet aircraft only.

The filtered dataset in saturation satisfies the following conditions: The number of jets taxiing out is greater than or equal to the number of 17, arrival throughput is less than or equal to 14, and jet departures are greater than or equal to 5. Figure 2-12 shows the jet departure throughput as a function of the arrival throughput and prop departures. We note that for this runway configuration at BOS, increasing the number of props in the fleet mix does not significantly decrease the departure throughput of jets. In particular, the curve of the average jet throughput given 3 prop takeoffs is



(a) Departure throughput as a function of arrival throughput and props departures using ASDE-X data for BOS configuration (VMC; 22L, 27 | 22L, 22R).

(b) Departure throughput as a function of arrival throughput and props departures using ASDE-X data for BOS configuration (VMC; 4R, 4L | 4R, 4L, 9).

Figure 2-11: BOS operational throughput envelopes estimated with ASDE-X data.

one unit lower than that of the average jet throughput given 0 props. This means that on average, for every three prop departures, one fewer jet will takeoff. The total departure throughput of the airport will increase by 2 AC/15 min in agreement with the plots of Figure 2-7.

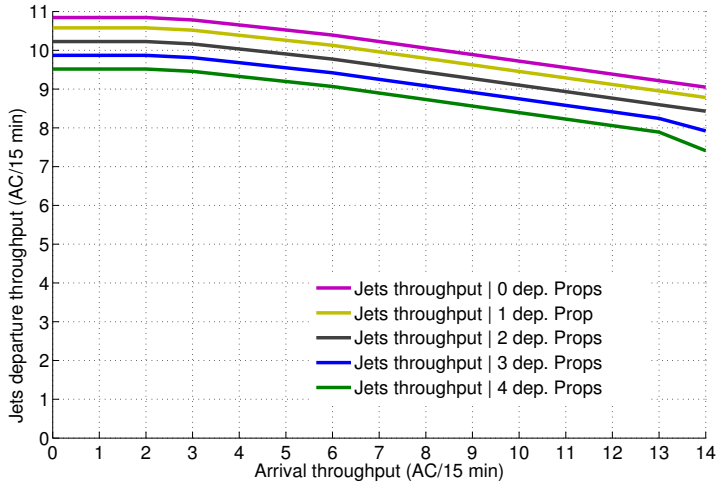


Figure 2-12: BOS departure throughput tradeoff between props and jets in configuration (VMC; 22L, 27 | 22L, 22R).

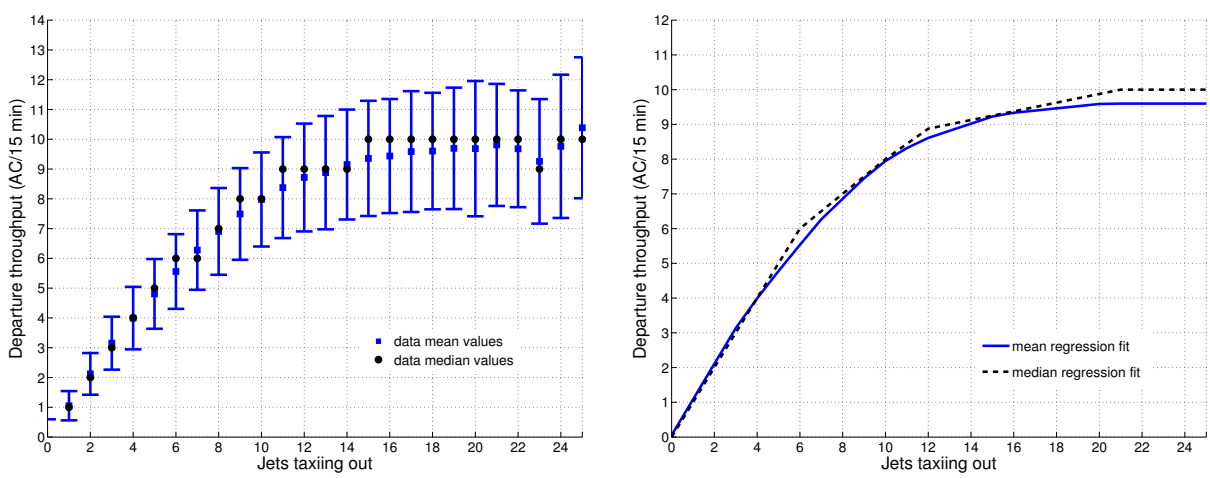
Figures 2-7 and 2-12 demonstrate that there is little interaction between jets and props for this particular runway configuration at BOS. For small numbers of props typically seen in the departure mix (the mean value of props departures is 1.04 and the median 1/15 min), the reduction of the jet departure throughput is at most 0.5 operations. This observation validates the hypothesis of

the air traffic controllers that prop departures have a small impact on jet departure throughput. Next, we study the jet departure throughput as a function of jet departure demand and arrivals, neglecting the impact of prop departures.

2.6.2 Saturation plot for jet departures

For this section, we use the whole dataset excluding all prop departures. As a first step, we generate a saturation plot, similar to the analysis in Section 2.1. The jet departure demand, N_J, at some time t is defined as the number of jets taxiing out at that time. The jet departure throughput during a 15-minute period starting at time t is defined as the jet departure throughput, T_J, over this time period and is measured as the number of jets that took off during the 15-minute interval $[t, t + 15)$.

This representation yields the plots of Figure 2-13a, which show the mean and median jet departure throughput for each value of the jet departure demand, N_J. The error bars depict one standard deviation of the departure throughput at each value of N_J. In Figure 2-13b, we show the corresponding regression fits. The mean jet departure throughput saturates at 9.6 AC/15 min when $N_J \geq 20$ and median jet departure throughput saturates at 10 AC/15 min when $N_J \geq 21$. We observe that the average jet departure throughput for this runway configuration under persistent demand is around 10 AC/15 min. The fitted curves suggest that persistent demand is achieved when the number of jets taxiing out exceeds 20.



(a) Measurements of the jet departure throughput as function of jets taxiing out.

(b) Regression of the jet departure throughput as function of jets taxiing out.

Figure 2-13: BOS jet departure throughput in configuration (VMC; 22L, 27 | 22L, 22R).

2.6.3 Estimation of the saturation point

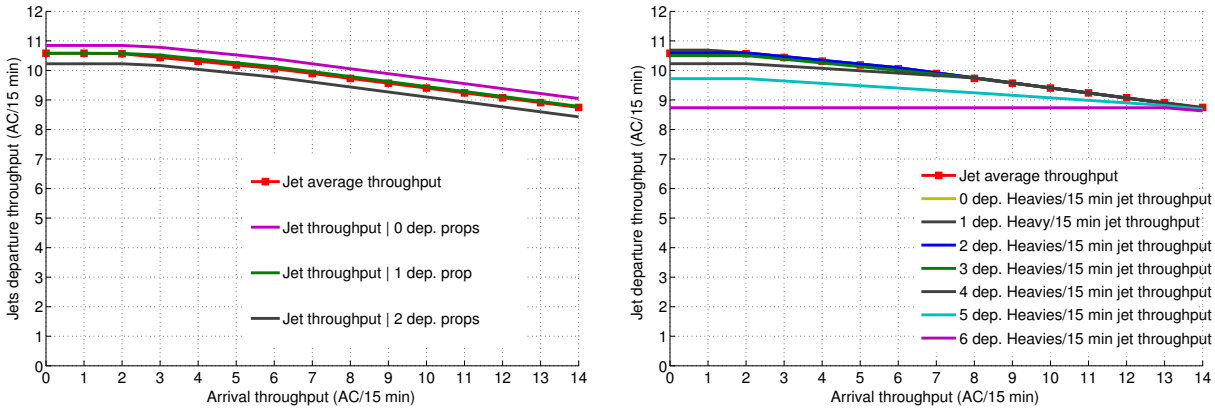
While the mean and median throughput measurements and fitted curves coincided in Figure 2-1 during congestion, they are different in Figure 2-13. This discrepancy is a result of excluding props. The median estimates are more robust to outliers, such as an unusually high number of props, which would decrease the number of jet operations even under high demand. The fitted curves also saturate at a very high value of N_J, that is, at $N_J = 21 = N^* - 1$, also as a result of excluding props. By contrast, the regression tree method estimates that $N_J^* = 17$.

2.6.4 Jet departure throughput as a function of arrival throughput and fleet mix

In a similar manner to Section 2.3, we isolate the filtered dataset in saturation and estimate the average jet departure throughput as a non-increasing, concave function of the arrival throughput. The fitted curve can be seen in Figure 2-14a in red, and is labeled *Jet average throughput*. We also plot three curves from Figure 2-12, namely the average jet throughput given 0, 1 and 2 prop departures. As a reminder, these throughput plots are fitted curves of the jet departure throughput, parametrized by the number of prop departures. We note that the *Jet average throughput* curve almost coincides with the curve of *Jet throughput | 1 dep. prop* and lies between the curves of *Jet throughput | 0 dep. props* and *Jet throughput | 2 dep. props*. This observations suggests that estimating the jet departure throughput while neglecting prop operations does not bias the estimation, but it reflects the average mix between jets and props.

We study the impact of Heavy departures on jet departure throughput. In Section 2.4, it was demonstrated that Heavy departures are a less significant explanatory variable of the departure throughput than the number of prop departures and the arrival throughput are. Now, we visualize this result by estimating the jet departure throughput as a function of the arrival throughput and the number of Heavy departures in a 15-minute interval. The graph of the estimated function is shown in Figure 2-14b. The jet departure throughput is insensitive to the number of Heavy departures, as long as the number of Heavies is not higher than 3, which is true during 95% of the high demand periods at BOS. We also note that the departure throughput when 5 Heavies depart is at most 9.7 jet/15 min, which is consistent with two minute separation requirement after a Heavy departure. Similarly, when the number of Heavies is 6, the jet departure throughput is around 9, as expected. Finally, it is not surprising that the throughput curve given six Heavy departures

is flat. The large number of Heavies results in many long separation requirements during which practically all arrivals can cross the departure runway, making departure throughput independent of the number of arrivals.



(a) Jet departure throughput as a function of arrival throughput and departing props.

(b) Jet departure throughput as a function of arrival throughput and Heavy departures.

Figure 2-14: Operational jet throughput envelopes in configuration (VMC; 22L, 27 | 22L, 22R).

2.7 Conclusions

This chapter introduced new methods for the parametric estimation of the departure capacity. We showed that the departure throughput can be estimated and represented as a function of departure demand and arrival throughput. We also showed how to measure departure capacity as a function of the arrival throughput and the fleet mix.

For the case of the 22L, 27 | 22L, 22R configuration of BOS, we showed that the average departure capacity (11 AC/15 min) can range from 8.5 to 14.5 AC/ 15 min, depending on arrival throughput, and fleet mix. We demonstrated that for this runway configuration, the fleet mix is a more significant explanatory variable of the departure throughput than the arrival throughput is, and compared our results to those of the state-of-the-art capacity envelope estimation technique. The comparison suggested that the differences in results were primarily because previous approaches do not consider the effect of fleet mix.

We also presented a methodology for studying the interactions between different types of aircraft. This analysis indicated that jet operations are decoupled from prop operations at BOS. Moreover, it showed that jet throughput is not adversely impacted by the Heavy jets among the departing aircraft, for reasonable numbers of Heavy aircraft.

In Chapter 3 and in Appendices D and H, we apply the methods introduced in this chapter to the analysis of capacity in different environments, such as, Newark Liberty International Airport, La Guardia Airport, the New York Metroplex, the major departure runway at Philadelphia International Airport, an individual runway at Dallas/ Fort Worth airport, and the new runway capacity at Charlotte Douglas International Airport.

Chapter 3

Case Study: Capacity Analysis of the New York Metroplex

In this chapter, we analyze the operational throughput of major airports in the New York (NY) Metroplex.

We first estimate the capacity of the Newark Liberty Airport (EWR) airport, quantifying the tradeoffs between departure and arrival throughput in different configurations. We then investigate the impact of the fleet mix of both arriving and departing aircraft on the departure throughput of EWR. After a detailed analysis of these parameters, we propose a compact and robust characterization of the capacity of the airport. We also identify the impact of the surrounding airspace conditions on the operational throughput of EWR. In order to do this, we investigate potential tradeoffs between EWR operations and those at the other major airports in the New York Metroplex, namely, John F. Kennedy International Airport (JFK), and LaGuardia Airport (LGA). Subsequently, we estimate the capacity of the airport system consisting of these three airports and discuss opportunities for improvement.

Finally, we introduce a recently developed tool, the Route Availability Planning Tool (RAPT) [32], and explain how it can be used to estimate the operational throughput of LGA in the presence of convective weather. In particular, we measure the effect of available airspace capacity on the operational throughput of the airport.

The following analysis builds on the work of Donaldson and Hansman [33], who provide a detailed description of the current operations in New York Metroplex, identify opportunities for improvement, and formulate hypotheses regarding the sources of inefficiency associated with current

operations. We use ASPM data of 2007 for all estimation problems unless otherwise mentioned.

3.1 Operational throughput of EWR

In this section, we adapt the methodology developed in Chapter 2 to the capacity estimation of the three major runway configurations of EWR, which are listed in Table 3.1. The saturation curves look very similar to those in Figure 2-1 and can be found in earlier work [106, 112]. EWR is a very congested airport and reaches more congested states than what was observed at BOS.

Table 3.1: Frequency of use of most frequently used EWR configurations under Visual and Instrument Meteorological Conditions in 2007 and 2008.

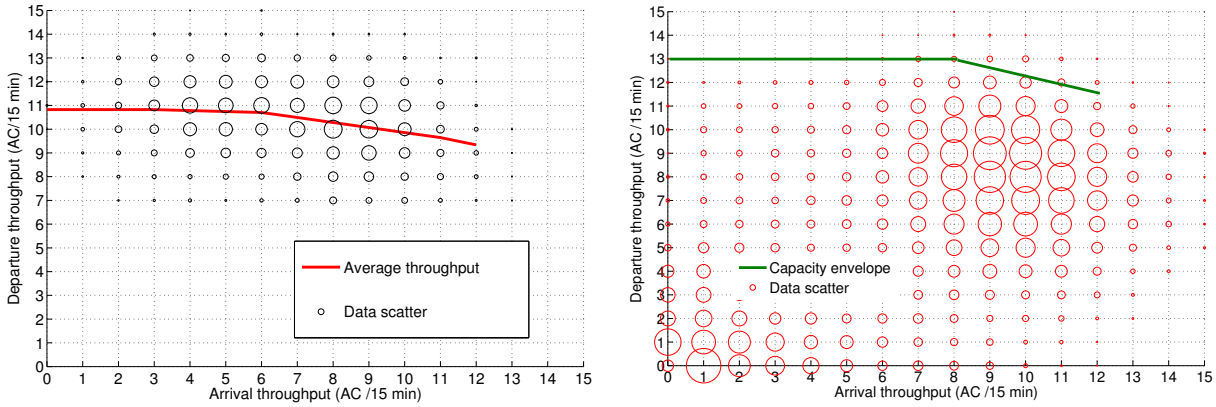
Runway Configuration	VMC		IMC	
Arrival Runway(s) | Departure Runway	Count (hrs)	Frequency	Count (hrs)	Frequency
22L | 22R	4453	40.3%	831	40.0%
11, 22L | 22R	2033	18.4%	74	3.6%
4R | 4L	2369	21.5%	981	47.2%

3.1.1 Estimation of operational throughput for configuration (VMC; 4R | 4L)

Following the method outlined in Section 2.3, we extract the filtered dataset in saturation, and use its data-points to estimate the average departure throughput as a non-increasing, concave function of the arrival throughput.

The scatterplot of the datapoints to which the fitting is applied, and the plot of the estimated function (in red) are shown in Figure 3-1a. The plot of Figure 3-1a provides the operational throughput envelope for this runway configuration at EWR. For example, for an arrival rate of 10 AC/15 min, the average departure throughput of this runway configuration under persistent departure demand and runway availability is 10 AC/15 min. This datapoint implies a balanced capacity of 20 movements/15 min, or 80 movements/hour. These values are close to the FAA estimates which place the capacity of this runway configuration at 80-81 movements per hour [83, 41].

Similarly to the BOS study, we compare the estimated function plotted in Figure 3-1a to traditional empirical capacity envelope estimates, shown in Figure 3-1b. The maximum total capacity is achieved at around 23-24 movements/15 min for several combinations of departure and arrival counts. This number would translate to 92-96 operations per hour, which is significantly higher than the number of movements that this runway configuration can sustain [41, 83].



(a) Operational throughput envelope.

(b) Capacity envelope.

Figure 3-1: Two different measures of the capacity of runway configuration (VMC; 4R | 4L) of EWR.

Impact of fleet mix on departure capacity

In this section, we quantify the impact of the fleet mix of both arriving and departing aircraft on departure capacity. According to the qualitative analysis of earlier work [112], our hypothesis is that the fleet mix can be represented with three variables in the case of EWR:

- Number of Heavy departures (H_{Deps}) in the 15-minute interval.
- Number of Heavy arrivals (H_{Arrs}) in the 15-minute interval.[1]
- Number of propeller-powered aircraft or props (P_{Deps}) departing in the 15-minute interval.

The mutual information scores between each potential explanatory random variable considered and the departure throughput are shown in Table 3.2, along with the correlation coefficients. We observe that the arrival throughput explains most of the variation in the departure throughput. Quite surprisingly, the Heavy arrivals appear to impact departure throughput to nearly the same extent as the Heavy departures. We finally note that the departing props appear to be an irrelevant variable in the case of this runway configuration of EWR.

Next, we consider a multi-variable model. We model the response variable, departure throughput (T), in each 15-minute time interval $[t, t + 15)$, as a function of five potential explanatory variables:

[1] In this analysis, we exclude B757s from the Heavy aircraft class, because B757s were shown to behave like Large aircraft with regard to departure throughput.

Table 3.2: Mutual information and correlation between departure throughput and potential explanatory variables.

Feature	Mutual Information with Departure throughput (T)	Correlation with Departure throughput (T)
Departure demand (N)	0.020	-0.033
Arrival throughput (A)	0.058	-0.243
Heavy departures (H_{Deps})	0.026	-0.148
Prop departures (P_{Deps})	0.001	0.015
Heavy arrivals (H_{Arrs})	0.019	-0.150

1. Number of departing aircraft (N) on the ground at time t. [2]

2. Number of arrivals (A) in the 15 minute interval.

3. Number of props taking off in the 15-minute time interval (P_{Deps}).

4. Number of Heavy departures in the 15-minute time interval (H_{Deps}).

5. Number of Heavy arrivals in the 15-minute time interval (H_{Arrs}).

We analyze the correlations between all the variables in the model using the *pairs* function of the R programming language [96], as explained in Section 2.4.2. The results are shown in Figure 3-2.

In Figure 3-2, we observe that the lowess fit line for the variable pair (N, T) does not exhibit any large or systematic deviation from the area between the $T = 10$ and $T = 11$ lines. The lack of dependence of the departure throughput on the departure demand validates the extraction of the *filtered dataset in saturation*. We note that the lowess fit line for the variable pair (A, T) follows the same trend as the operational throughput envelope (Figure 3-1a). The departure throughput decreases from 11 to 9 operations/15 min as a concave function of the arrival throughput.

The lowess fit line for the variable pair (H_{Deps}, T) decreases gradually from 11 to 10 AC/15 min as the number of Heavy departures increases from 0 to 8 AC/15 min. As was seen in the case of BOS the rate of decrease is lower than what is theoretically expected. Heavy aircraft departures require twice as long a wake vortex separation requirement as Large aircraft. However, it appears that wake vortex separation is not the major driver of the departure throughput of this runway configuration. The lowess fit line for the variable pair (H_{Arrs}, T) exhibits a clear negative correlation between the two variables. As the number of Heavy arrivals increases from 0 to 3, the

[2] Although we have established that in the filtered dataset in saturation the departure throughput does not change significantly with the number of aircraft on the ground it is useful to revisit this hypothesis in the multi-variable model.

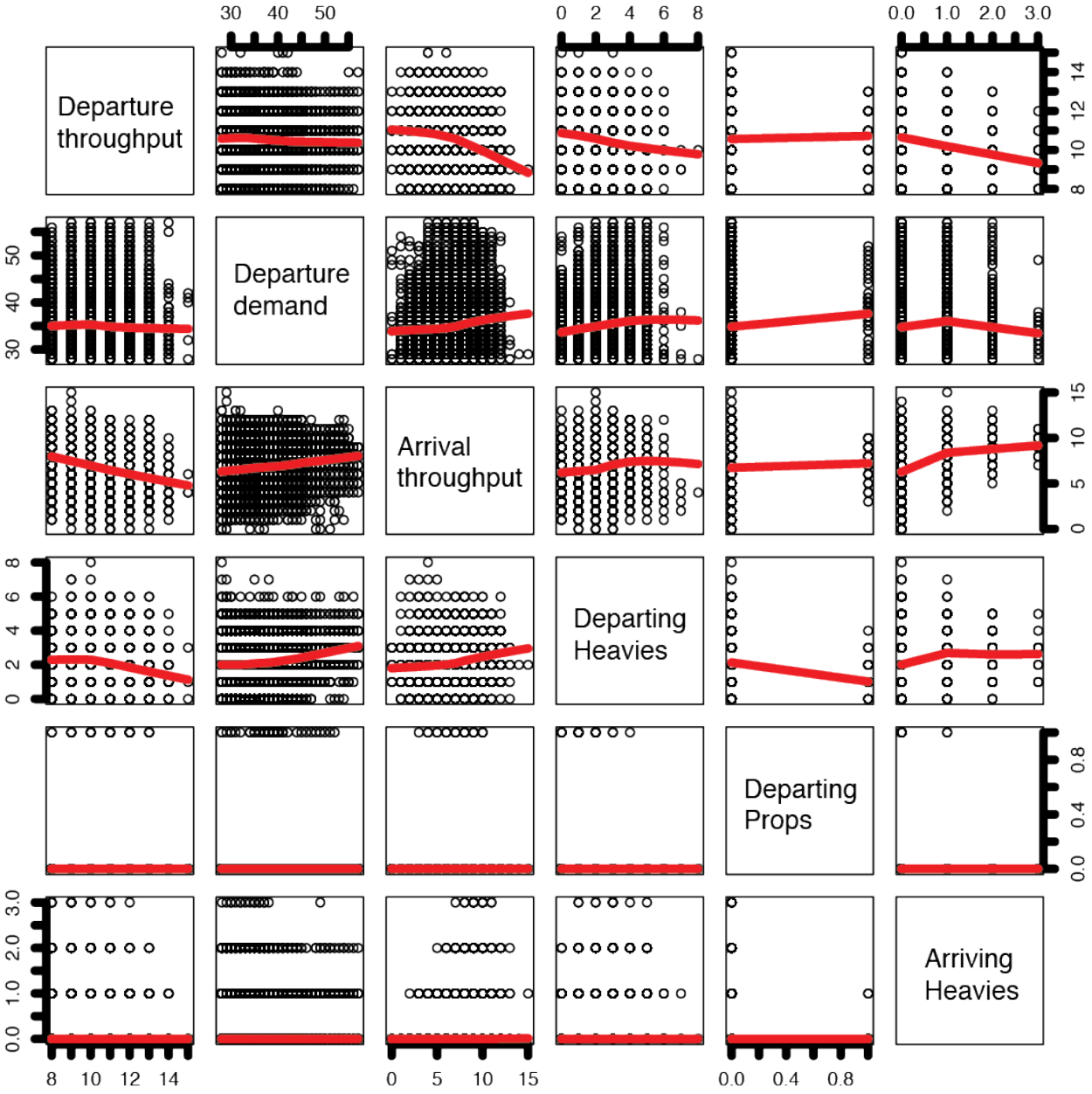


Figure 3-2: Correlation between departure throughput, departure demand, arrival throughput, Heavy departures, prop departures and Heavy arrivals.

departure throughput decreases from 11 to about 9. The lowess fit line for the variable pair (P_{Deps}, T) does not show any clear relation between the two.

As was done for BOS, we examine the relationship between the departure throughput and these five variables with other statistical tools as well (regression trees, random forests and generalized additive models). They all lead to the same conclusion that the three most significant explanatory variables are the number of arrivals, and the number of Heavy aircraft (among both arrivals and

departures).

A tradeoff between Heavy arrivals and total departure throughput was not expected because of the segregation of arrivals and departures in the configurations analyzed. However, Table 3.2 and the plots of Figure 3-2 show that, in fact, Heavy arrivals do have a significant negative impact on the departure throughput. Possible causes of this tradeoff are either that Heavy arrivals request the departure runway, which is 1000 ft longer than the arrival runway, or the additional time required for Heavy arrivals to cross the departure runway due to slower taxi acceleration. As a result, air traffic controllers either have Heavy aircraft exit the arrival runway and immediately cross the departure runway without slowing down, or bring the Heavy aircraft to a full stop before crossing the departure runway. Both maneuvers disrupt the departure flow.

This conjecture also explains why the impact of Heavy arrivals is diluted when classifying arriving B757s as *Heavy arrivals*. The taxi characteristics of B757s are similar to those of Large aircraft, and they do not typically request a longer runway. By contrast, departing B757s have a longer wake vortex separation requirement, similar to Heavy departures.

Parametric representation of departure capacity

Having established that the departure throughput is primarily a function of arrival throughput, Heavy arrivals and Heavy departures, we estimate and plot this function. This is a challenging task because (1) it is hard to visualize a three variable function, and (2) the impact of the explanatory variables can be coupled. For these reasons, we adopt an approximate approach: We first estimate the departure throughput in the filtered dataset in saturation as a function of the arrival throughput and the number of Heavy departures in a 15-minute interval, neglecting the impact of Heavy arrivals.

We follow a similar procedure to that described in Chapter 2: Given k triplets of measurements $A(t)$, $H_{Deps}(t)$ and $T(t)$, denoted by $(u_1, v_1, y_1), \ldots, (u_k, v_k, y_k)$, we seek a function $g_{h1} : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $T(t) = g_{h1}(A(t), H_{Deps}(t))$. The constraints are imposed only between neighboring points, as was explained in Section 2.4.3. The problem is formulated as follows, where $l = \max(A(t))$ and $n = \max(H_{Deps}(t))$:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{3.1}$$

subject to:

$$\hat{y}_i = g_{h1}(u_i, v_i), \;\; i = 1, \ldots, k \tag{3.2}$$

$$g_{h1}(i+1, j) \leq g_{h1}(i, j), \;\; i = 0, \ldots (l-1), \forall j \tag{3.3}$$

$$g_{h1}(i+1, j) - g_{h1}(i, j) \leq g_{h1}(i, j) - g_{h1}(i-1, j), \;\; i = 1, \ldots (l-1), \forall j \tag{3.4}$$

$$g_{h1}(i, j+1) \leq g_{h1}(i, j), \;\; j = 0, \ldots (n-1), \forall i \tag{3.5}$$

$$g_{h1}(i, j+1) - g_{h1}(i, j) \leq g_{h1}(i, j) - g_{h1}(i, j-1), \;\; j = 1, \ldots (n-1), \forall i \tag{3.6}$$

$$g_{h1}(i, j+1) - g_{h1}(i, j) \leq g_{h1}(i+1, j+1) - g_{h1}(i+1, j), \;\; i = 0, \ldots (l-1), j = 0, \ldots (n-1) \tag{3.7}$$

$$g_{h1}(i, j) - g_{h1}(i+1, j) \geq g_{h1}(i, j+1) - g_{h1}(i+1, j+1), \;\; i = 0, \ldots (l-1), j = 0, \ldots (n-1) \tag{3.8}$$

Here, Inequalities (3.3) and (3.4) are analogous to those in the case of the capacity envelope, namely, for a given fleet mix, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. Inequality (3.5) ensures that for fixed arrival throughput, the departure throughput is a non-increasing, concave function of the number of Heavy departures. This relation holds because each Heavy aircraft departure introduces a two minute separation requirement. Heavy aircraft departures are also expected to result in increasing capacity loss as their number increases: For a small number of Heavies, the extra separation requirement that they introduce can be used to cross arrivals or sequence traffic. However, as the number of Heavy aircraft increases, the wake vortex separation requirements become a tight constraint: For seven Heavies, at most eight departures can be accommodated in that 15-min interval.

Similarly, Equation (3.7) ensures that the marginal loss in departure throughput from increasing the number of Heavy departures by one unit decreases as the arrival throughput increases. This is because there is more pressure from arriving aircraft to cross runway 4L as the number of arrivals increases, and this pressure reduces the impact of Heavy departures. The departure runway will be utilized for crossing arrivals independent of the wake vortex separation requirements of departing aircraft.

Finally, Equation (3.8) ensures that the marginal loss in departure throughput from increasing the arrival throughput by one unit increases as the number of departing Heavies increases. If increasing the arrival throughput by one unit forces the airport to decrease departure throughput by some amount, then decreasing the arrival throughput by one unit and replacing a Large aircraft with a Heavy one will result in at least the same decrease in departure throughput.

The plot of the estimated function, $g_{h1}(A, H_{Deps})$, is shown in Figure 3-3a overlaid with the red curve of Figure 3-1b (the *average throughput*). The comparison shows that the lines in Figure 3-3a are in fact the red line parametrized by the number of Heavy aircraft departing in the 15-minute interval.

We also estimate the departure throughput as a function of the arrival throughput and the number of Heavy arrivals in a 15-minute interval neglecting the impact of departing Heavies.

Given k triplets of measurements $A(t)$, $H_{Arrs}(t)$ and $T(t)$, denoted by $(u_1, w_1, y_1), \ldots, (u_k, w_k, y_k)$ in the filtered dataset in saturation, we seek a function $g_{h2} : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $T = g_{h2}(A(t), H_{Arrs}(t))$. The problem formulation is as follows, where $l = \max(A(t))$ and $n = \max(H_{Arrs}(t))$:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{3.9}$$

subject to:

$$\hat{y}_i = g_{h2}(u_i, w_i), \;\; i = 1, \ldots, k \tag{3.10}$$

$$g_{h2}(i+1, j) \leq g_{h2}(i, j), \;\; i = 0, \ldots (l-1), \forall j \tag{3.11}$$

$$g_{h2}(i+1, j) - g_{h2}(i, j) \leq g_{h2}(i, j) - g_{h2}(i-1, j), \;\; i = 1, \ldots (l-1), \forall j \tag{3.12}$$

$$g_{h2}(i, j+1) \leq g_{h2}(i, j), \;\; j = 0, \ldots (n-1), \forall i \tag{3.13}$$

$$g_{h2}(i, j+1) - g_{h2}(i, j) \geq g_{h2}(i, j) - g_{h2}(i, j-1), \;\; j = 1, \ldots (n-1), \forall i \tag{3.14}$$

$$g_{h2}(i, j+1) - g_{h2}(i, j) \leq g_{h2}(i+1, j+1) - g_{h2}(i+1, j), \;\; i = 0, \ldots (l-1), j = 0, \ldots (n-1) \tag{3.15}$$

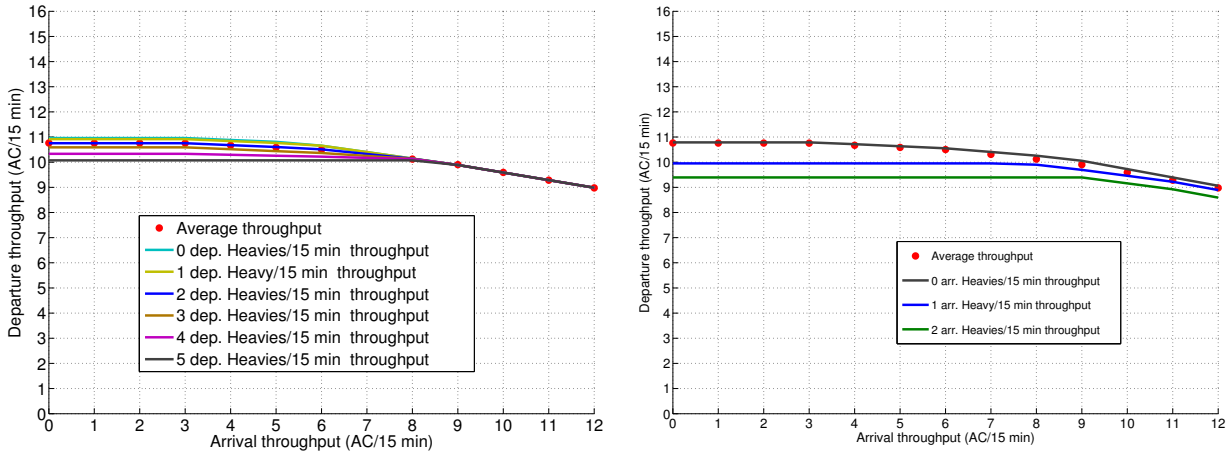
$$g_{h2}(i, j) - g_{h2}(i+1, j) \geq g_{h2}(i, j+1) - g_{h2}(i+1, j+1), \;\; i = 0, \ldots (l-1), j = 0, \ldots (n-1) \tag{3.16}$$

Inequalities (3.11)- (3.12) are analogous to the ones in the case of the capacity envelope. For a given fleet mix, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. Inequality (3.13) ensures that for fixed arrival throughput, the departure throughput is a non-increasing function of the number of Heavy arrivals. This relation holds because Heavy arrivals are expected to slow departure throughput at least as much as non-Heavy arrivals. Inequality (3.14) ensures that the departure throughput is a convex function of the number of Heavy arrivals, for a given total number of arrivals. The rationale behind this is that replacing

Small or Large aircraft with Heavy ones, will lead to diminishing capacity losses, since multiple crossings of Heavy aircraft can be combined.

Similarly, Equation (3.15) ensures that the marginal loss in departure throughput from increasing the number of Heavy aircraft by one unit decreases as the number of arrivals increases. Finally, Equation (3.16) ensures that the marginal loss in departure throughput from increasing the arrival throughput by one unit increases as the number of Heavy arrivals increases. If increasing the arrival throughput by one unit results in a certain decrease of the departure throughput, then increasing the arrival throughput by one unit and replacing a Large arrival with a Heavy one will yield at least the same decrease in departure throughput.

The plot of the estimated function, $g_{h2}(A, H_{Arrs})$, can be seen in Figure 3-3b overlaid with the red curve of Figure 3-1b (*average throughput*).



(a) Departure throughput as function of the arrival throughput and the number of Heavy departures.

(b) Departure throughput as a function of the arrival throughput and the number of Heavy arrivals.

Figure 3-3: Parametric representation of the capacity of the 4R|4L runway configuration.

With regard to Figure 3-3a we first observe that the average departure throughput curve coincides with the curve describing the departure throughput conditioned on two Heavy departures in a 15-minute interval. This is consistent with the proportion of Heavy aircraft in the fleet mix at EWR, which was around 22% in 2007. Similar to the BOS results (Section 2.6.4), the number of departing Heavy aircraft does not significantly impact the departure throughput. For the most common operating scenarios in which the rate of arrivals is 5-10 AC/15 min and the number of Heavy departures is 1-3, the departure throughput does not change substantially. As the number of arrivals increases, the impact of Heavy departures diminishes, and all the curves in Figure 3-3a

converge to a single one, similar to what was observed for BOS.

By contrast, the number of Heavy arrivals has a higher impact on departure throughput than Heavy departures, when measured in terms of throughput reduction/ Heavy operation. For the most common operating scenarios in which the rate of arrivals is 5-10 AC/15 min, two Heavy arrivals reduce the departure throughput by 0.5-1 operation per 15-min (i.e., 2-4 operations per hour). As the number of arrivals increases, the impact of Heavy arrivals decreases from more than 1 fewer departure per 2 Heavy arrivals, to less than half a departure loss per two Heavy arrivals. We hypothesize that controllers use the extra time required to cross a Heavy arrival to perform multiple crossings in periods with many landings. In this manner, the impact of Heavy arrivals diminishes. A comparison of the curves in Figures 3-3a and 3-3b suggest that Heavy arrivals are more detrimental to departure throughput than Heavy departures. Two Heavy arrivals can introduce a departure throughput penalty of one operation per 15 minutes, whereas two Heavy departures do not reduce the departure efficiency.

We also observe that for low numbers of arrivals and no Heavy aircraft in the fleet mix, the departure capacity is less than 11 AC/15-min. In theory, one would expect this number to be closer to 15AC/15-min. The standard terminal radar separation between non-departing Heavies translates to approximately one minute separation between successive takeoffs with the same heading. The fact that it is significantly lower than the theoretical estimate suggests that other constraints (for example, TRACON capacity, En Route Center capacity, or traffic flow management programs) decrease the operational departure throughput of this runway configuration.

Finally, we note that such a parametric representation assumes that there is no correlation between Heavy departures and Heavy arrivals; an analysis of the relation between these variables is presented in Section 3.1.4.

3.1.2 Analysis of South flow configurations

In this section, we apply the methodology presented in section 3.1.1 to the two other major runway configurations at EWR, namely, 22L | 22R and 11, 22L | 22R. These configurations are primarily used under south winds.

As explained by Donaldson and Hansman [33], Runway 11 at EWR is used under south winds and high arrival demand. However, several restrictions apply due to its short length (6800 ft), and restricted approach geometry. Its use is typically limited to Boeing 737-700 and smaller aircraft and 15-mile spacing is required between successive arrivals. It is worth noting that the use of Runway

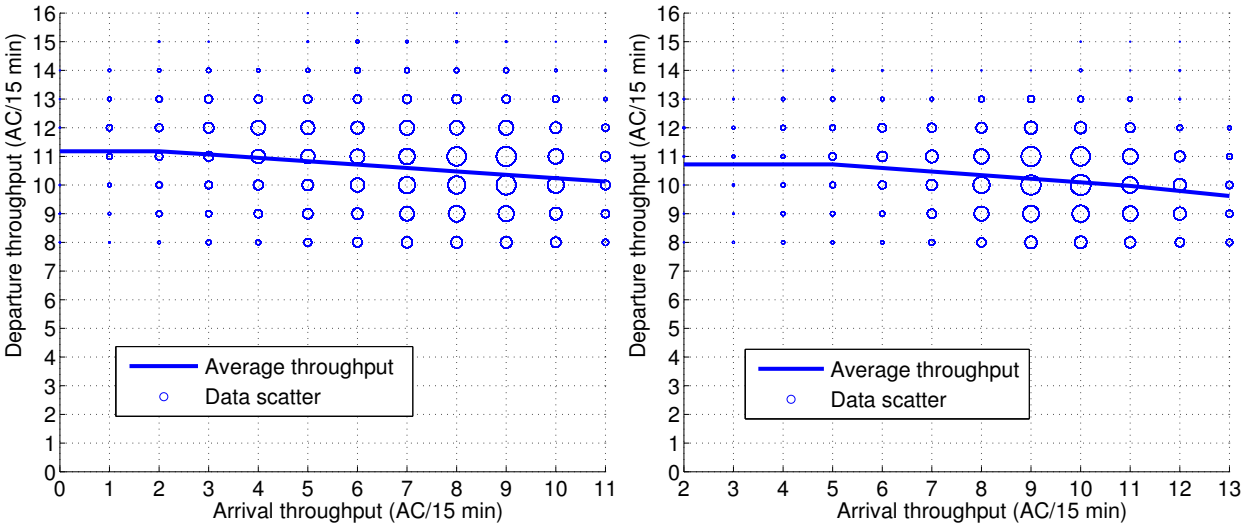
11 often happens with a slight tailwind component, as the southerly prevailing winds are between 210° and 270° . Above a certain threshold, this tailwind component can lead to an increased miles–in–trail (MIT) restriction of 20 miles between arrivals to Runway 11. Departures on 22R can be operated independently from the two arrival runways as long as takeoffs begin south of the intersection with Runway 11 [33]. The subsequent analysis in this chapter will identify the extent of added arrival capacity offered by Runway 11, and its impact on the departure capacity.

Runway configuration 22L | 22R is symmetric to 4R | 4L, so one would expect the two configurations to have the same capacity. However, the airspace design differs substantially between the two runway configurations: Airspace to the south is less restricted than that to the north, allowing aircraft to be sent to two different (dispersal) departure headings (215° and 239°) immediately after takeoff from runway 22R. This procedure means the inter-departure spacing requirement is reduced to 6000 ft or when the leading departure becomes airborne (unless the leading aircraft is a Heavy or B757). Without multiple dispersal headings, subsequent departures from Runway 4L must be given sufficient spacing to ensure that the required 2.5 mile terminal radar separation can be maintained between them [33]. This typically translates to a one-minute separation requirement. Our analysis investigates whether runway configuration 22L | 22R has a higher departure capacity than 4R | 4L, as a result of its less restricted airspace.

Operational throughput envelopes

Following the method outlined in Section 2.3, we isolate the filtered dataset in saturation, and use its data-points to estimate the average departure throughput as a concave non-increasing function of the arrival throughput. The resulting operational throughput envelopes are shown in Figure 3-4.

Comparing Figures 3-1a, 3-4a and 3-4b , we first observe that runway configuration 11, 22L | 22R tends to be used under high arrival demand. There are no data-points with fewer than 2 arrivals/ 15 min, and most of the data-points are concentrated in high arrival counts, that is between 9 and 12 arrivals. By contrast, the data-points for runway configuration 22L | 22R are spread across a broad range of arrival counts. With regards to the impact of Runway 11 on departure capacity, we note that for arrival counts that can be accommodated by both configurations (4-12 arrivals/ 15 min), the departure capacity of the two configurations is not significantly different. This suggests that the landings on Runway 11 do not impact the departures from 22R. We also note that the addition of Runway 11 increases the arrival priority capacity by only two additional landings. Such a small improvement may be explained by the MIT restrictions that apply to the use of Runway 11,



(a) Operational throughput envelope of configuration (VMC; 22L | 22R).

(b) Operational throughput envelope of configuration (VMC; 11, 22L | 22R).

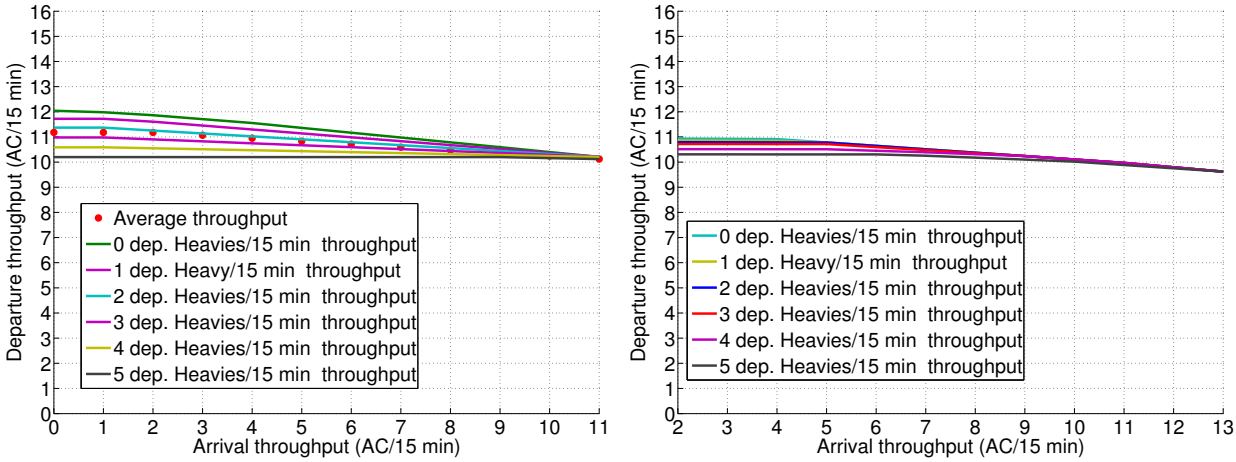
Figure 3-4: Operational throughput envelopes of the two major south-flow runway configurations at EWR.

and the fact that landings on it need to be sequenced with landings on 22L. Indeed, according to the FAA EWR Traffic Management Tips [41], the AAR of runway 22L is 28-42 arrivals/hr, while that of Runway 11 is only 4-6 arrivals/hr. The throughput envelopes show that the departure capacity of runway configurations 22L | 22R and 11,22L | 22R is not significantly higher than that of the 4R | 4L, in agreement with the FAA EWR Traffic Management Tips [41]. This finding suggests that, on average, dispersal headings do not significantly increase departure capacity, possibly because of fleet mix and taxiing-in aircraft crossing the departure runway.

Impact of departing Heavy aircraft

Following the procedure described in Section 3.1.1, we estimate the impact of Heavy departures on the departure throughput for the 22L | 22R and 11, 22L | 22R runway configurations. The results of the estimation procedure are shown in Figure 3-5.

From Figure 3-5a, we note that the departure throughput of the 22L | 22R runway configuration increases from 11 AC/15 min to 12 AC/15 min, given no Heavy departures and a small number of arrivals. This is in contrast to the departure throughput of the 4R | 4L configuration, which does not increase substantially with a non-Heavy fleet mix (Figure 3-3a). We hypothesize that the availability of dispersal headings for takeoffs from runway 22R explains this difference. From Figure



(a) Departure throughput as a function of the arrival throughput and the number of Heavy departures for configuration (VMC; 22L | 22R).

(b) Departure throughput as a function of the arrival throughput and the number of Heavy departures for configuration (VMC; 11, 22L | 22R).

Figure 3-5: Impact of Heavy departures on operational throughput for the two major south-flow runway configurations at EWR.

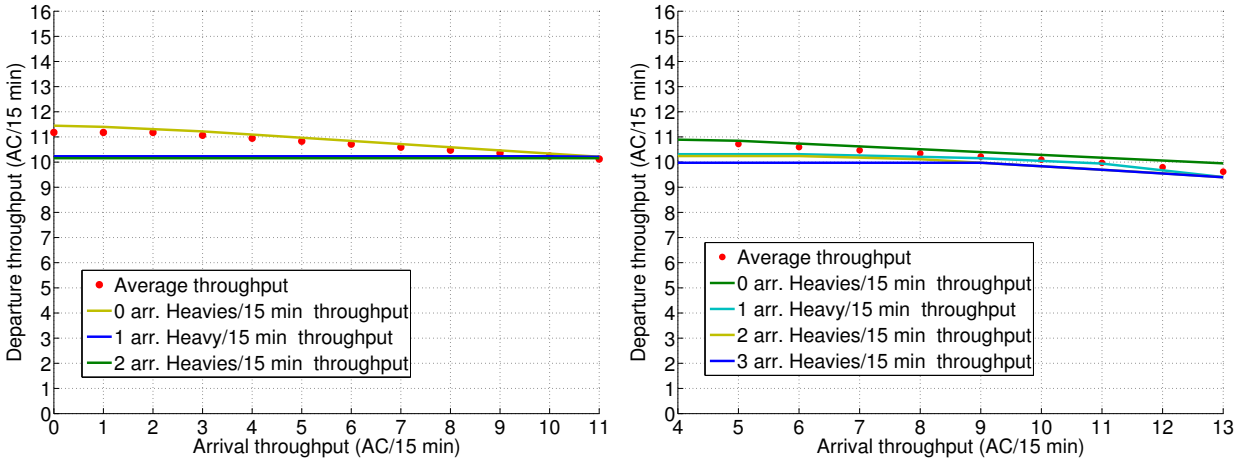
3-5b we observe that the departure throughput of the 11, 22L | 22R configuration does not vary significantly with the number of Heavy aircraft in the departing fleet mix. This can be explained by the high arrival rate served by this runway configuration since air traffic controllers most probably use the large separation requirement behind a Heavy departure to have arrivals cross the departure runway. For the same reason, the throughput curves of Figure 3-5a converge as the number of arrivals increases for the 22L | 22R configuration as well.

Impact of Heavy arrivals

Here, we estimate the impact of Heavy aircraft arrivals on the south–flow runway configurations. The results can be seen in Figure 3-6. Heavy arrivals impact both south-flow runway configurations in a similar fashion in the common range of arrival rates (4-12 arrivals/ 15 min). In the range of 4-8 arrivals/15 min, two Heavy arrivals come at approximately the cost of half a takeoff. The effect of Heavy arrivals diminishes as the total number of arrivals increases, similar to the impact of Heavy arrivals on departure throughput for the 4R | 4L configuration (Figure 3-3b).

3.1.3 Predictive capabilities of EWR models

Similarly to the analysis of BOS departure throughput in Section 2.5, we provide five measures for goodness-of-fit of the proposed estimation as applied to EWR. We compare four different estimates



(a) Departure throughput as a function of arrival throughput and the number of Heavy arrivals for configuration (VMC; 22L | 22R).

(b) Departure throughput as a function of arrival throughput and the number of Heavy arrivals for configuration (VMC; 11, 22L | 22R).

Figure 3-6: Impact of Heavy aircraft arrivals for the two major south–flow runway configurations at EWR.

for $\hat{y}_i$:

1. The arithmetic mean (AM) of the departure throughput : $y = \frac{1}{k}\sum_{i=1}^{k} y_i$.

2. The departure throughput predicted from the operational throughput envelope (OTE): $h(u_i)$.

3. The predicted throughput from the operational throughput envelope parametrized for Heavy departures (PTE1): $g_{h1}(u_i, v_i)$.

4. The predicted throughput from the operational throughput envelope parametrized for Heavy arrivals (PTE2): $g_{h2}(u_i, w_i)$.

Table 3.3 lists aggregate statistics for the dataset which is used for estimating the throughput curves for the three major runway configurations at EWR. Comparing them to the corresponding ones for BOS, shown in Table 2.2, we notice that the mean departure throughput of the major runway configuration at BOS, 22L, 27 | 22L, 22R, in 2007 (10.56) is approximately the same as the mean departure throughput of the major runway configuration of EWR, 22L | 22R, in 2007 (10.50). By contrast, the variance of the departure throughput of the major runway configuration at BOS (3.72) is much higher than the variance of the departure throughput of the major runway configuration of EWR (2.25). We also notice that the dataset in the case of EWR is much larger than that of BOS, because EWR is a much more congested airport. This results in more data-points in the filtered dataset in saturation.

Table 3.3: Statistics of the training dataset for the EWR analysis.

Runway Configuration	Time frame	Source	Size	Mean	Variance
4R | 4L	Jan-Dec 2007	ASPM	10076	10.52	1.70
22L | 22R	Jan-Dec 2007	ASPM	18926	10.50	2.25
11, 22L | 22R	Jan-Dec 2007	ASPM	8622	10.19	1.73

Table 3.4 lists the goodness of fit results for the four estimators. In contrast to the case of BOS, the parametrized capacity envelopes explain very little of the variation of the departure throughput. There are two reasons for this: Firstly, as already discussed, the variance of the departure throughput of EWR is much smaller. Secondly, in contrast to BOS, none of the potential explanatory variables explain a significant portion of the variation. From Figures 3-3 to 3-6, we note that in all of the parametrized throughput envelopes and for most of the ranges of the explanatory variables considered, the departure throughput is within one operation of the 10 AC/ 15 min point.

Table 3.4: Comparison of four different estimators for the departure throughput on the EWR training data set.

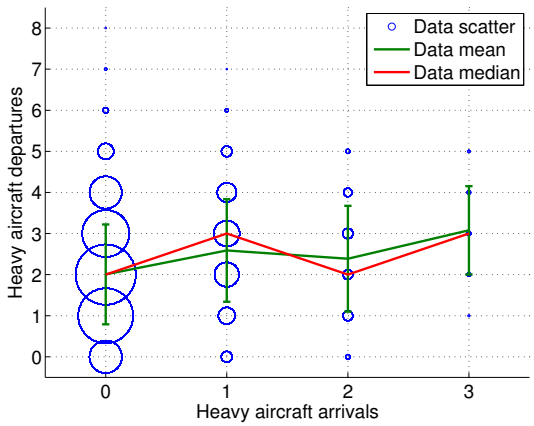
	Estimator	MAE	MSE	RMSE	MAPE	RMSPE
4R | 4L	AM	1.08	1.70	1.30	10.6%	13.2%
4R | 4L	OTE	1.02	1.59	1.26	10.0 %	12.7%
4R | 4L	PTE1	1.00	1.56	1.25	10.0%	12.6%
4R | 4L	PTE2	1.01	1.58	1.26	10.0%	12.7%
22L | 22R	AM	1.25	2.25	1.500	1.22 %	15.0%
22L | 22R	OTE	1.19	2.15	1.47	11.7%	14.6%
22L | 22R	PTE1	1.17	2.10	1.45	11.5 %	14.5%
22L | 22R	PTE2	1.17	2.11	1.45	11.5%	14.6%
11, 22L | 22R	AM	1.08	1.73	1.31	10.8%	13.4%
11, 22L | 22R	OTE	1.05	1.67	1.29	10.6%	13.2%
11, 22L | 22R	PTE1	1.05	1.66	1.29	10.6%	13.2%
11, 22L | 22R	PTE2	1.04	1.64	1.29	10.5 %	13.1%

3.1.4 Correlation between Heavy departures and Heavy arrivals

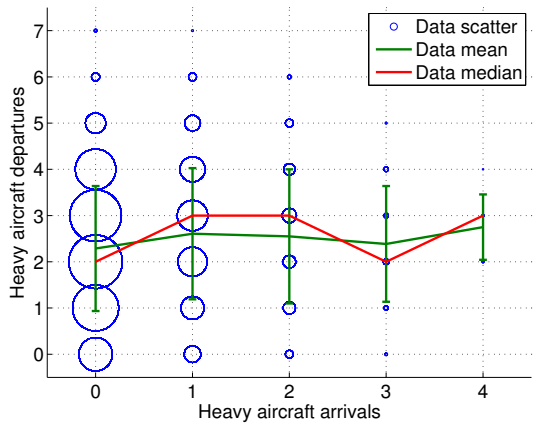
In Sections 3.1.1 and 3.1.2, we separately quantified the impact of Heavy arrivals and Heavy departures by estimating the departure throughput, first as a function of arrival throughput and Heavy departures, and then as a function of arrival throughput and Heavy arrivals. In other words, the curves of Figure 3-3a regress the departure rate on arrival throughput and Heavy departures neglecting the impact of Heavy arrivals. For interpreting the results of this regression, it is essential

to investigate potential correlation between the Heavy departures and the omitted variable, Heavy arrivals. For example, if zero Heavy departures are highly correlated with zero Heavy arrivals, the curves "0 dep. Heavies/15 min throughput" of Figure 3-3a and "0 arr. Heavies/15 min throughput" of Figure 3-3b would estimate the same quantity, namely, the departure throughput as a function of the arrival count, conditioned on zero Heavy departures and zero Heavy arrivals.

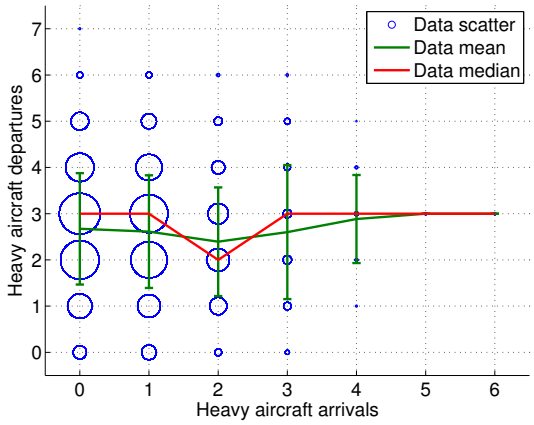
Next, we test for the correlation between the two variables, *Heavy departures* and *Heavy arrivals*. In Figure 3-7, we plot the number of Heavy departures and the number of Heavy arrivals as a scatter plot for the three major runway configurations at EWR.



(a) Heavy Departures as a function of Heavy arrivals at runway configuration 4R | 4L.



(b) Heavy Departures as a function of Heavy arrivals at runway configuration 22L | 22R.



(c) Heavy Departures as a function of Heavy arrivals at runway configuration 11, 22L | 22R.

Figure 3-7: Relation between Heavy departures and Heavy arrivals for three major runway configurations at EWR.

We also show the mean and median values of the number of Heavy departures, conditioned on

the number of Heavy arrivals. The error bars represent one standard deviation of the distribution. From these plots, we first note that there is no clear trend between Heavy departures and Heavy arrivals. At all values of Heavy arrivals, there are, on average, 2-3 Heavy departures. We also note that the variance of Heavy departures does not change across the range of Heavy arrivals with sufficient data-points, that is, 0-3 Heavy arrivals. For all values of Heavy departures, the most data-points correspond to periods with no Heavy arrivals. This implies that the curves of Figure 3-3a and Figure 3-5, in which the Heavy arrivals are neglected, are representative of no Heavy arrivals, because most of the observations are under this condition.

Similarly, the regression fits of Figure 3-3b and Figure 3-6 neglect the impact of Heavy departures. Since there is no correlation between Heavy departures and Heavy arrivals, neglecting the Heavy departures does not introduce an omitted-variable bias in the fits. In addition, for all values of Heavy arrivals, the observed counts of Heavy departures are more more evenly spread in the range of 1-4 per 15 min. The fits of Figure 3-3b and Figure 3-6 are therefore representative over a range of values of Heavy departures. Table 3.5 also presents the correlation and the mutual information between the two variables for the three major configurations.

Table 3.5: Correlation between Heavy departures and Heavy arrivals for the prominent runway configurations at EWR.

Runway configuration	Correlation between Heavy departures and Heavy arrivals	Mutual information Heavy departures and Heavy arrivals
4R | 4L	0.176	0.031
22L | 22R	0.157	0.026
11,22L | 22R	-0.013	0.009

3.1.5 Concluding remarks on EWR analysis

Studying the three prominent VMC runway configurations at EWR, we conclude that the departure throughput is inelastic to changes in arrival throughput. It is around 10 AC/15 min for all runway configurations, and varies only by 1 AC/15 min when the arrival throughput takes values in the range of 0-11 AC/15 min.

In Table 3.6, we summarize our findings for the three different configurations. Under the column "Balanced Operations", we present the capacity of the airport when it serves arrivals and departures in equal numbers. Its capacity is approximately 20 AC/ 15 min, or 80 AC/hour. In the column "Arrival Priority", we present the capacity of the airport during periods in which the

highest values of arrival throughput are realized. In this regime, the capacity is approximately 12 Arrivals/15 min and 9 Departures/15 min, or 48 Arrivals/hour and 36 Departures/hour. Similarly, in the column "Departure Priority", we present the capacity of the airport during periods in which the highest values of departure throughput are realized. In this case, the capacity is approximately 6 Arrivals/15 min and 11 Departures/15 min, or 24 Arrivals/hour and 44 Departures/hour.

We notice that merely capping operations at 81 Movements/hour, like the slot control policy proposed by FAA in 2009 [36], may not be very effective in demand management and delay reduction, since the allocation of movements to departures and arrivals also needs to be specified. For example, there is no operational regime in which the airport can sustain either 81 Departures/hour or 81 Arrivals/hour, as demonstrated by Pyrgiotis [93]. Therefore, a more effective slot control policy would be to cap operations in a manner similar to Table 3.6, instead of using a single number.

Table 3.6: EWR aggregate average runway throughput (AC/15 minutes).

Capacity	Balanced Operations	Arrival Priority	Departure Priority
Arrival	10	12	6
Departure	10	9	11
Total	20	21	17

3.2 Capacity of the New York Metroplex

It has been conjectured that the potential capacity of EWR is not achieved due to interactions with the other airports in the New York Metroplex [33]. In section 3.1, we saw that the departure capacity is relatively insensitive to variables that theoretical models suggest as being very influential, such as, the departing aircraft fleet mix. The results of Section 3.1 motivate an investigation of EWR's interactions with the other two major airports in the New York Metroplex, JFK and LGA.

We first investigate whether these interactions explain some of the variation in EWR's departure capacity. Subsequently, we examine tradeoffs between the operations of the three major airports in New York. We conclude this section by characterizing the capacity of the airport system comprising JFK, EWR and LGA.

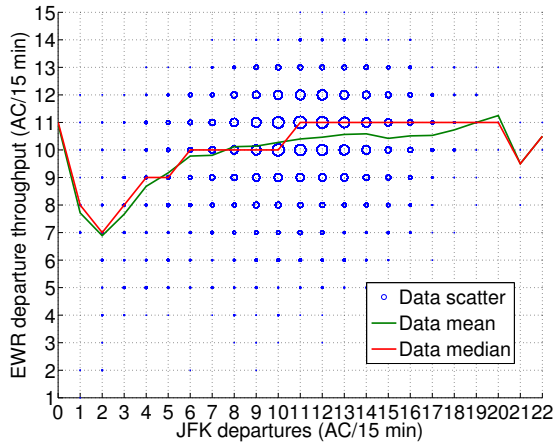
3.2.1 Interactions between JFK and EWR

As a first step, we study the impact of JFK departures on EWR departures, to investigate potential tradeoffs between the two airports. This tradeoff could be caused because of merging traffic at departure gates, or by TRACON workload constraints [33]. In Figure 3-8, we show the scatter plot of the departure throughput of EWR in saturation (i.e., under persistent demand)[3] and the departure throughput of JFK, for the three major runway configurations at EWR. We also show the mean and median trend of the EWR departure throughput in saturation, conditioned on the number of departures at JFK.

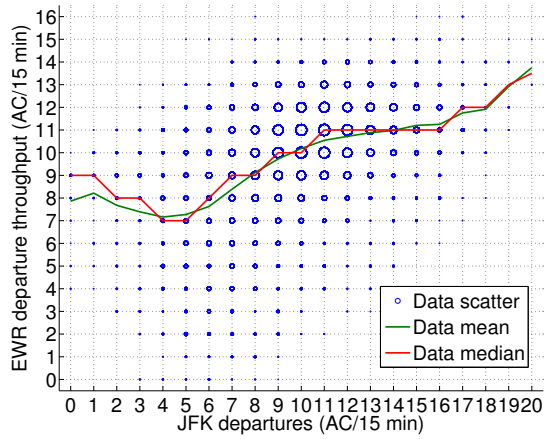
From Figure 3-8, we observe that there is no trend indicating that JFK departures negatively impact EWR departure throughput for any of the EWR configurations. In fact, the plots suggest a positive trend, which is most pronounced for the 22L | 22R configuration. The more efficient JFK is, the more efficient EWR is as well. This positive correlation can be explained by the airspace design, which keeps the departure flows out of the two airports sufficiently separated (Figure B-1). Departures from JFK therefore do not negatively impact those from EWR. By comparing Figures 3-8a and 3-8b, we note that the mean EWR departure throughput does not exceed 11 AC/15 min for all values of JFK departures in the 4R | 4L configuration. However, in the 22L | 22R configuration, the mean departure throughput at EWR exceeds 11 AC/15 min for high values of JFK departures. This suggests that when JFK departure throughput is very high, EWR throughput also increases, in the 22L | 22R configuration. High JFK departure throughput implies high route availability in the New York airspace, which, in turn, is correlated with high route availability for EWR departures. Because the airspace to the south of EWR is not constrained, departures make use of the increased routing options, resulting in increased departure throughput. However, when the airport operates in the 4R | 4L configuration, the departure throughput cannot increase as much because of the airspace constraints.

In order to further investigate the correlation between the departure capacities of the two airports, we extract the data-points for which JFK is under persistent demand (in saturation). The hypothesis we want to test is whether high JFK demand results in its prioritization. In Figure 3-9, we show the scatter plot of the departure throughput of EWR in saturation and the departure throughput of JFK in saturation for the three major runway configurations at EWR. We also show the mean and median trends of EWR departure throughput in saturation, conditioned on the

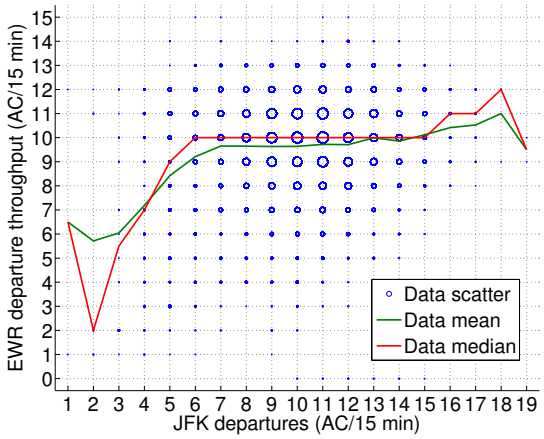
[3] In this section we do not filter low throughput values from the dataset in saturation, because we test the hypothesis of whether they are related to operations at other airports of the NY Metroplex.



(a) EWR departure capacity as a function of JFK departures for the 4R | 4L configuration at EWR.



(b) EWR departure capacity as a function of JFK departures for the 22L | 22R configuration at EWR.

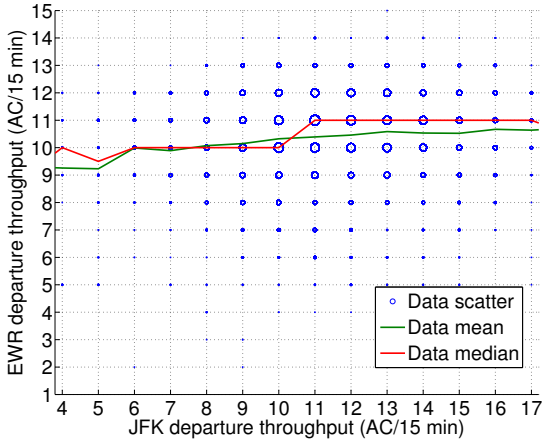


(c) EWR departure capacity as a function of JFK departures for the 11,22L | 22R configuration at EWR.

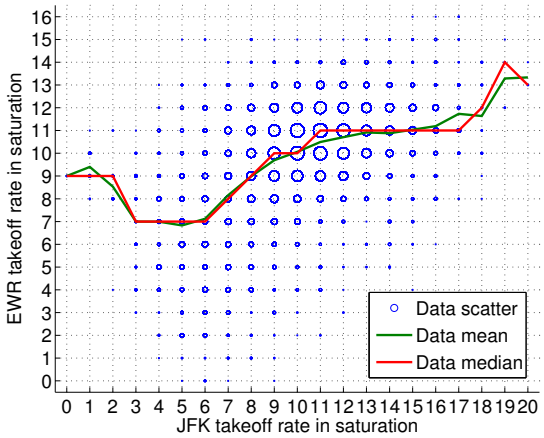
Figure 3-8: Relation between EWR departure capacity and JFK departures for the three major runway configurations at EWR, during periods when EWR is in saturation.

departure throughput of JFK in saturation.

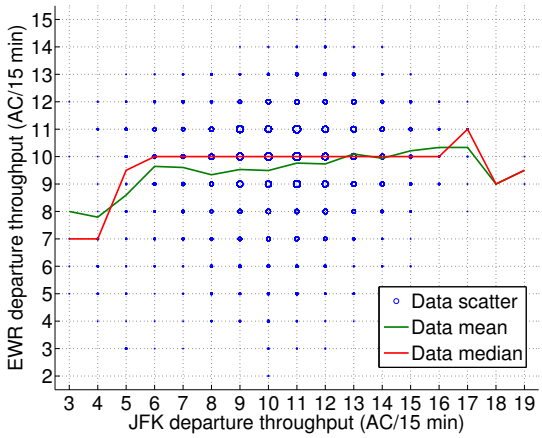
The plots of Figure 3-8 and Figure 3-9 are quite similar, suggesting that JFK departures interact with EWR departures in the same manner irrespective of whether or not JFK is saturated. The positive correlation between EWR and JFK capacity suggests the presence of hidden variables which impact the departure capacity of the two airports in a similar fashion. One such factor could be the arrival demand: When one airport experiences an arrival bank, the other airport is likely to experience one as well. Similarly, both airports experience a large number of Heavy departures in the evening. Another factor could be downstream weather, or route availability. We note that when



(a) EWR departure capacity as a function of JFK departure capacity for the 4R | 4L configuration at EWR.



(b) EWR departure capacity as a function of JFK departure capacity for the 22L | 22R configuration at EWR.



(c) EWR departure capacity as a function of JFK departure capacity for the 11,22L | 22R configuration at EWR.

Figure 3-9: Relation between EWR departure capacity and JFK departure capacity, for three major runway configurations at EWR during periods when both airports are in saturation.

the departure throughput of JFK is very low (fewer than 7 AC/15 min), the departure throughput of EWR is low as well. In periods when both airports are under persistent departure demand and both have very low departure throughput, it is likely that downstream constraints keep both airports at low performance levels.

To further emphasize the correlation between EWR and JFK departure throughputs under persistent demand, we compare the correlation coefficient between the two variables to the correlation coefficient between EWR departure throughput in saturation and EWR arrival throughput (Table

3.7). We note the absolute values of the correlation between the departure capacities of the two airports is higher than those between the departure capacity of EWR and the arrival capacity of EWR, for two out of the three major configurations. This preliminary analysis therefore shows no evidence that JFK departures negatively impact EWR departures.

Table 3.7: Correlation between EWR departure capacity and JFK departure capacity.

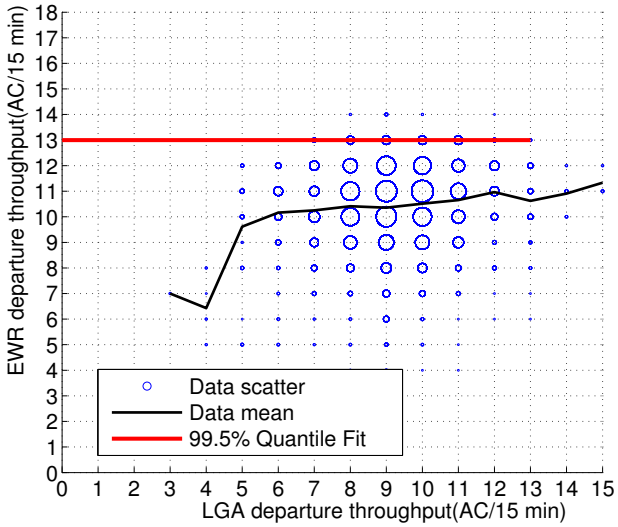
EWR runway configuration	Correlation between EWR dep. capacity and JFK dep. capacity	Correlation between EWR dep. capacity and EWR arr. capacity
4R | 4L	0.142	-0.261
22L | 22R	0.474	-0.053
11, 22L | 22R	0.138	-0.082

3.2.2 Interactions between LGA and EWR

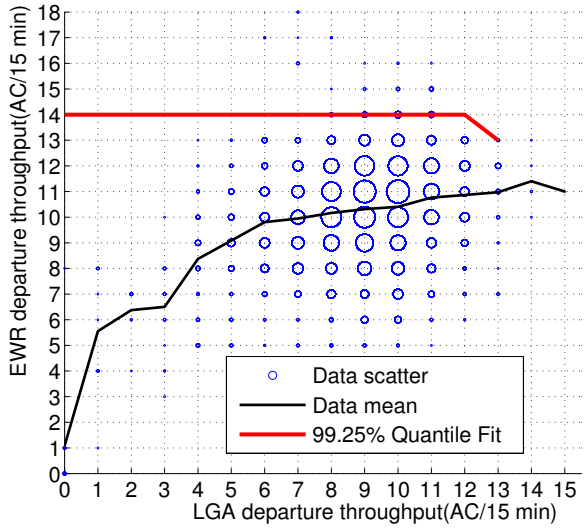
In this section, we investigate potential departure tradeoffs between EWR and LGA. In contrast to the previous section, we look at configurations that are frequently in use simultaneously at the two airports. Since JFK is hypothesized to be prioritized over the other New York airports, our intention in Section 3.2.1 was to study how departures out of EWR are impacted by different levels of departure throughput at JFK. Here, we focus on configuration combinations of EWR and LGA that are known to share airspace resources, in particular departure fixes [33]. These configurations are listed in Table 3.8.

Table 3.8: Frequency of simultaneous use of EWR and LGA configurations under Visual Meteorological Conditions.

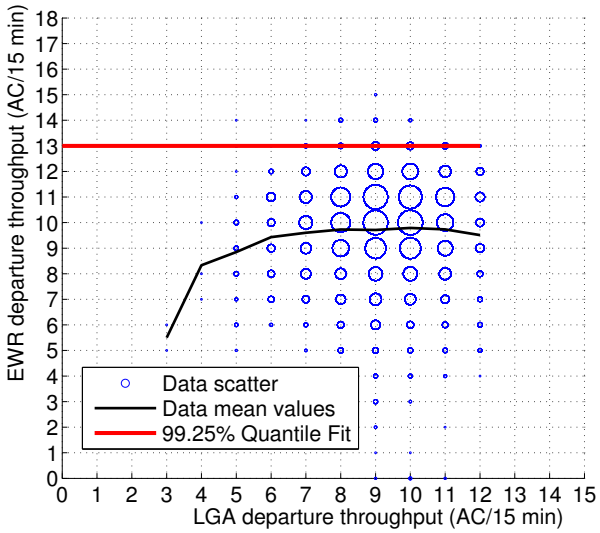
Airport(s)	Runway Configuration	Use	
		Count (hrs)	Frequency
EWR	22L | 22R	2664	36.7 %
EWR	4R | 4L	1519	20.9%
EWR	11, 22L | 22R	1398	19.25 %
LGA	31 | 4	1400	19.2 %
LGA	22 | 31	1567	21.6 %
LGA	22 | 13	16585	22.8 %
(EWR, LGA)	(22L | 22R; 31 | 4)	874	12.4 %
(EWR, LGA)	(4R | 4L ; 22 | 31)	583	8.3 %
(EWR, LGA)	(11, 22L | 22R ; 22 | 13)	583	8.3 %



(a) EWR departure capacity as a function of LGA departure capacity for the (22L | 22R; 31 | 4) configuration.



(b) EWR departure capacity as a function of LGA departure capacity for the (4R | 4L ; 22 | 31) configuration.



(c) EWR departure capacity as a function of LGA departure capacity for the (11, 22L | 22R ; 22 | 13) configuration.

Figure 3-10: Relation between EWR departure capacity and LGA departure capacity for the three major runway configurations at EWR and LGA.

In Figure 3-10, we show the scatter plot of the departure throughputs of the two airports in saturation (i.e., under persistent departure demand). We also show the average EWR departure throughput as a function of the LGA departure throughput. We note that similar to JFK, LGA departure throughput does not appear to have negative correlation with the average EWR departure throughput. For all runway configuration combinations, the average departure throughput of EWR

increases with increasing departure throughput of LGA.

Figure 3-10 also shows the quantile regression fits. The quantiles are chosen according to the methodology suggested by Ramanujam and Balakrishnan [97]. The quantile regression could be informative if LGA had an adverse impact on the departure throughput of EWR, but this could not be discerned by studying the average performance of EWR as a function of LGA departure throughput. Average departure throughput is a multi-variable function, and the effect of LGA alone cannot be isolated. However, by studying the statistically significant maximum observed number of operations with quantile regression, we can find out whether increasing LGA throughput reduces the maximum throughput achievable by EWR. The quantile regression is in agreement with our previous observations, and in agreement with the results of Ramanujam and Balakrishnan [97]. The departure throughputs of the two airports do not interact, and EWR can attain its maximum throughput independent of LGA departure throughput. We also note that the maximum throughput of EWR is higher under the 22L | 22R configuration because of the higher departure route availability, as discussed in Section 3.1.2.

3.2.3 Operational throughputs of JFK and LGA

For completeness, we provide the operational throughput envelopes for the major runway configurations of JFK and LGA in Figures C-1 and C-2 of Appendix C. In Tables 3.9 and 3.10, we summarize capacity estimates of these two airports. Comparing Tables 3.6, 3.9 and 3.10, we note that the three major New York airports have very similar balanced operations capacity values (20 AC/15 min for JFK and EWR, and 18 AC/15 min for LGA). We also note that the departure priority capacity is inelastic in all airports. Finally, we note that runway configuration 31L, 31R | 31L at JFK, which favors arrival throughput, is the only runway configuration that exhibits a clear tradeoff between arrival and departure throughput. Inspecting Figure C-1a, we note that in order to reach the arrival priority operating point, the departure throughput decreases to 5. In this runway configuration, the only departure runway, 31L, is shared by departures and arrivals.

Table 3.9: JFK aggregate average runway throughput (AC/15 min).

Capacity	Balanced Operations	Arrival Priority	Departure Priority
Arrival	10	16	6
Departure	10	6	12
Total	20	22	18

Table 3.10: LGA aggregate average runway throughput (AC/15 min).

Capacity	Balanced Operations	Arrival Priority	Departure Priority
Arrival	9	12	6
Departure	9	8	10
Total	18	20	16

3.2.4 Local regression for Metroplex

Building on the work of Donaldson and Hansman [33], we treat the whole Metroplex as a single airport system. Donaldson and Hansman identified three major Metroplex configurations, the most frequently-used of which is labeled "South-Flow-VMC-Arrival Priority". Its components can be seen in Table 3.11. For all individual runway configurations comprising "South-Flow-VMC-Arrival Priority", we calculate the saturation point using the methodology of Section 3.1.1. We now keep all data points for which the Metroplex is in "South-Flow-VMC-Arrival Priority" and all individual airports are in saturation.

Table 3.11: Elements of the "South-Flow-VMC-Arrival Priority" Metroplex configuration.

Airport	Runway configuration	Weather	N^*
JFK	13L, 22L | 13R	VMC	25
EWR	11, 22L | 22R	VMC	25
LGA	22 | 13	VMC	13

We study the interactions among the following variables:

1. JFK departure throughput: Departure throughput in each 15-min time interval at JFK.
2. JFK departure throughput: Arrival throughput in each 15-min interval at JFK.
3. EWR departure throughput: Departure throughput in each 15-min time interval at EWR.
4. EWR arrival throughput: Arrival throughput in each 15-min interval at EWR.
5. LGA departure throughput: Departure throughput in each 15-min interval at LGA.
6. LGA arrival throughput: Arrival throughput in each 15-min interval at LGA.

Similarly to Section 3.1.1, we analyze the correlations between all the variables using the *pairs* function of the R programming language, which produces panels with the correlations among all variables of the model. Each panel of Figure 3-11 shows the scatterplot between the variable on the vertical axis and the variable on the horizontal axis, as well as the lowess curve (in red)

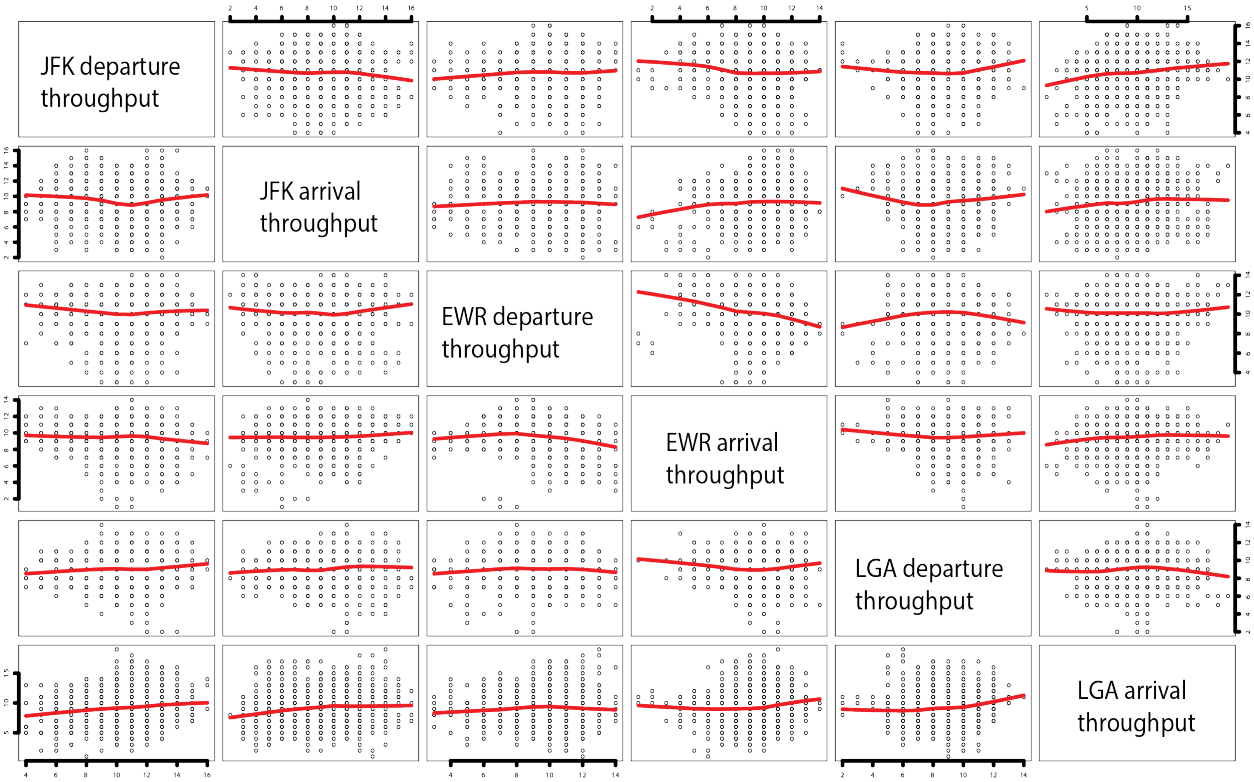


Figure 3-11: Correlations between departures and arrivals in the NY Metroplex for the "South-Flow-VMC-Arrival Priority" configuration.

through the set of data points. From the lowess fit lines, we observe that there is no clear negative correlation between departures out of one airport in the Metroplex, and operations at another. On the contrary, we note the decreasing trend between the departure throughput and the arrival throughput resulting from the shared capacity between departures and arrivals at each individual airport (the capacity envelope). We also note the non-decreasing trend between the departure throughput of two individual airports. This observation is consistent with the analysis of Section 3.2.1, which showed a positive correlation between the departure capacities of EWR and JFK. The trend between departures at LGA and departures at EWR is very similar to that in Figure 3-10c reinforcing the conclusion that LGA departures have little impact on EWR departures. Similarly, there are no decreasing trends between the departure throughput at one airport and the arrival throughput at another, or between arrival throughputs at different airports.

These results were verified using other statistical models, as well, such as generalized additive models, regression trees, and random forests. In summary, the analysis of the Metroplex configuration "South-Flow-VMC-Arrival Priority" shows that the departure and arrival capacities of an airport are not adversely impacted by operations at other airports.

3.2.5 Metroplex operational throughput

We characterize the capacity of the NY Metroplex for the "South-Flow-VMC-Arrival Priority" and two other major Metroplex configurations, as identified by Donaldson and Hansman [33], and shown in Tables 3.12 and 3.13.

Table 3.12: Elements of the "South-Flow-VMC-Departure Priority" Metroplex configuration.

Airport	Runway Configuration	Weather	N^*
JFK	22L | 22R, 31L	VMC	29
EWR	22L | 22R	VMC	26
LGA	22 | 31	VMC	13

Table 3.13: Elements of the "North-Flow-VMC-Departure Priority" Metroplex configuration.

Airport	Runway Configuration	Weather	N^*
JFK	4R | 4L, 31L	VMC	22
EWR	4R | 4L	VMC	28
LGA	31 | 4	VMC	17

We isolate the datapoints for which all three airports are under persistent departure demand and estimate the operational throughput envelopes of the three major Metroplex configurations (Figures 3-12, 3-13a and 3-13b). Each "arrival throughput" point consists of the sum of the arrival throughputs of the three airports, and each "departure throughput" point, the sum of their departure throughputs. In other words, we treat the Metroplex as a single airport. The estimates are shown in Table 3.14.

Comparing Figures 3-12 and 3-13, we notice that for the Departure Priority configurations, the arrival throughput takes lower values and the departure priority throughput is higher. We also notice that the departure throughput in South Flow is higher than the departure throughput in the "North-Flow-Departure-Priority" configuration. Consistent with this, the departure throughput of the three individual airports is higher when they operate in the corresponding runway configurations. Despite the symmetry in the airfield design of the runway configurations in the "Departure-Priority-South-Flow" and "Departure-Priority-North-Flow" Metroplex configurations, the airspace design is significantly different for the two flows. South flow offers more departure fixes, allowing for a higher number of departure operations for a given number of arrivals. We note that the departure throughput stays steady for the most part in the "South-Flow-VMC-Arrival Priority" configuration. The reason for this behavior is that the departure throughput changes

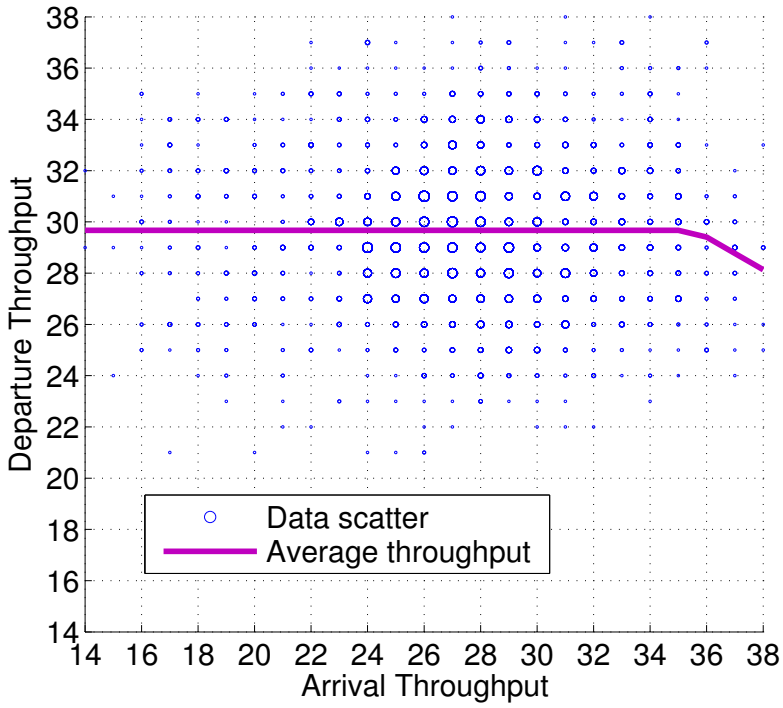


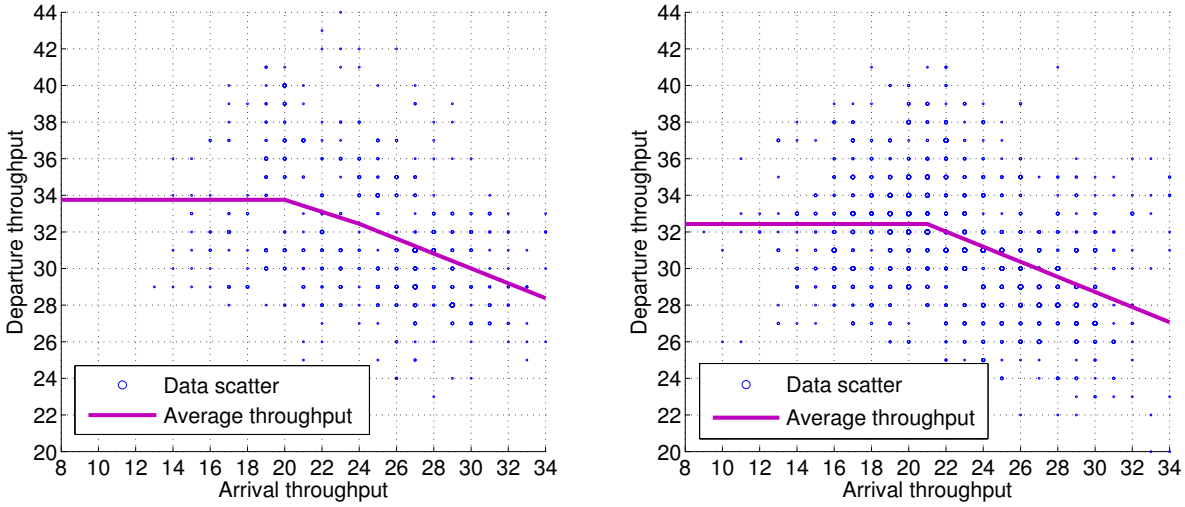
Figure 3-12: Operational throughput envelope of the New York Metroplex for the "South-Flow-VMC-Arrival Priority" configuration.

very little in each individual airport with arrival throughput: In the case of JFK, the departure throughput is in the range of 10-11 departures for 0-11 arrivals (Figure C-1a); in the case of EWR, 10-11 departures for 0-13 arrivals (Figure 3-4b); and in the case of LGA, 9 departures for 0-12 arrivals (Figure C-2a). Because of the averaging when treating the whole Metroplex as a single "airport", we observe a steady departure throughput of 30 aircraft for 14-35 arrivals. However, when the number of arrivals exceeds 35, the departure throughput starts decreasing.

Table 3.14: New York Metroplex aggregate average throughput (AC/15 minutes).

Capacity	Balanced Operations	Arrival Priority	Departure Priority
Arrival	30	38	20
Departure	29	28	33
Total	59	64	53

In conclusion, we note that even with the current airspace design of the New York Metroplex, there are opportunities for performance improvement. For instance, in the "South-Flow-VMC-Departure Priority" configuration, we saw that EWR does not appear to take advantage of the availability of dispersal headings. Similarly, in the "South-Flow-VMC-Arrival Priority" configuration, the added arrival runway increases the arrival priority capacity of JFK by only 3 AC/15



(a) Operational throughput envelope for the"South-Flow-VMC-Departure Priority" configuration.

(b) Operational throughput envelope for the"North-Flow-VMC-Departure Priority" configuration.

Figure 3-13: Operational throughput envelopes of the New York Metroplex for the Departure Priority configurations.

min.

Analogously, we note that JFK uses two departure runways in the Departure-Priority configurations, the operations of which are decoupled,[4] but yet achieves a departure throughput of only 13 AC/15 min (Figure C-1). For these configurations and considering the fleet mix of JFK, theoretical models predict much higher arrival and departure throughputs [33]. The investigation of the reasons behind the low throughput at JFK is out of the scope of this thesis, but is an interesting topic which, if addressed, could help improve airport congestion and resultant downstream delays without necessitating airspace redesign, new runways or other expensive capacity improvements. Our hypothesis is that aircraft are assigned to runways not with the objective of maximizing the airport departure throughput, but according to their preferred departure fixes. This practice is likely to result in frequent unbalanced utilization of the two departure runways, as has been seen to occur at other US airports, like MEM[5] and DFW[6].

[4]Departures start their takeoff roll on 31L past the intersection with 22R, thereby allowing independent operations on both runways.

[5]Interviews with the FedEx Surface Operations Management group.

[6]In Section D.2 of the Appendix, we discuss the impact of imbalanced runway utilization on the departure throughput at DFW.

3.3 Leveraging route availability information for operational throughput estimation

All applications of the proposed methodology demonstrated so far involved runway configurations during Visual Meteorological Conditions (VMC). Analogous plots can be constructed for Instrumental Meteorological Conditions (IMC). However, in addition to local meteorological conditions, downstream restrictions can also impact the operational throughput of an airport [60]. There may be thunderstorms blocking fixes out of the airport, or major departure routes. Aircraft that planned to use these fixes or routes would then have to be rerouted. As the number of blocked fixes or routes increases, the available rerouting options decrease, affecting aircraft in two ways: Some aircraft may need to be routed over a small number of fixes or even a single fix. The throughput of a single fix or route is typically much smaller than the capacity of a runway. The radar separation requirement is 5 nm, and multiple airports might make use of a single fix. In these cases miles–in–trial (MIT), minutes–in–trail (MINIT), or other traffic management initiatives be initiated at the origin airports to ensure that aircraft arrive at the over-loaded fix or route at a sustainable rate. Other flights, the available routes of which are blocked, will be delayed until one of their potential routes clears.

Both cases above decrease the number of the aircraft that are cleared for takeoff at a given time. As fewer and fewer aircraft are allowed to take off, the throughput of the airport is effectively decreased. Although capacity is theoretically available, it cannot be used because of downstream restrictions. The affected aircraft may be classified as "departure demand" by our methods, because in many cases, especially at the New York airports, aircraft are assigned route blockage-related delays after they have left their gate [116].

3.3.1 Route Availability Planning Tool (RAPT)

In this section, we discuss how information on downstream restrictions can be used for estimating the operational throughput of an airport. We focus on a particular class of downstream restrictions, namely, route blockage for departures out of an airport as measured with the Route Availability Planning Tool (RAPT). RAPT, developed by MIT Lincoln Laboratories, is an automated decision support tool intended to help air traffic controllers determine the specific departure routes and departure times that will be affected by operationally significant convective weather. RAPT helps users determine when departure routes or fixes should be opened or closed, and to identify alternative departure routes that are free of convective weather. It has been in use for the past 10 years

in the New York area airports and has reduced a great fraction of delays and costs associated with convective weather [31, 32, 100]. We use archived RAPT data from LGA in 2010-2011 to study the impact of route availability, as predicted by RAPT, on the capacity of the airport. We model the impact of route availability on the mean value and the variance of the throughput of the airport, in contrast to prior literature which focuses on calculating its impact on individual flights and fixes [31, 32, 100].

RAPT assigns one of three status colors/values–RED/3 (blocked), YELLOW/2 (impacted), DARK GREEN/1 (insignificant weather encountered) or GREEN/0 (clear)–to each route, every 5 minutes, for departure times up to 30 minutes into the future. In addition, it provides a wealth of auxiliary information, such as the location of the blockage and the echo-top (height) of the closest blockage on each route. However, we do not know how the RAPT status is actually used by controllers, and which aircraft are affected. Data on the preferred and actual routes flown by each aircraft are not available to us. We therefore use RAPT information in an aggregate way by averaging across both time and routes. For each route and each 5-minute period, we calculate a moving average of the seven entries of the "Blockage Status" of this route: the 15-min, 10-min and 5-min prior trends, the current time period trend, and the 5-min, 10-min and 15-min forecasts. The time-averaging is performed to model the use of the tool in practice, since controllers, when assessing the availability of a route, do not only consider its status in the current time, but its recent values and predicted future trend as well. The recent values are used in assessing the reliability and stability of the current trend. The future values are useful, as aircraft will use these routes in the future. We average across the 13 routes to obtain a single value for each 5-minute period to account for spatial errors. We discretize this value in steps of 0.2 and denote it as S_{RAPT} (Surrogate RAPT). In the absence of more detailed operational information, we believe that this simple metric is a good approximation of weather impact on departure routes.

3.3.2 Analysis of LGA with RAPT data

In LGA during 2011, runway configuration 22 | 13 was in use during most days in which departure routes were impacted by convective weather. Table 3.15 shows the statistics of use for this runway configuration for the years 2010-2011, while Table 3.16 shows the usage of 22 | 31 .

For estimating the operational throughput, we extend the methodology developed in Section 2.3.1 as follows: For each runway configuration, we obtain the data for which the departure throughput does not change significantly with the departure demand, that is, the data for which

Table 3.15: RAPT indicators when runway configuration 22 | 13 was in use at LGA.

S_{RAPT} value	Use (hrs) in 2010		Use (hrs) in 2011	
	Total	Saturation periods	Total	Saturation periods
0.0	1027	453	900	359
0.2	73	33	76	29
0.4	43	20	47	25
0.6	28	14	28	13
0.8	8	3	15	9
1	4	3	6	5
1.2	1	1	0	0
All	1184	527	1072	440

Table 3.16: RAPT indicators when runway configuration 22 | 31 was in use at LGA.

S_{RAPT} value	Use (hrs) in 2010		Use (hrs) in 2011	
	Total	Saturation periods	Total	Saturation periods
0.0	1073	160	1228	139
0.2	41	8	29	7
0.4	28	10	11	1
0.6	31	7	3	1
0.8	12	5	3	1
1	6	1	0	0
1.2	1	0	0	0
All	1192	191	1274	149

$N^* \leq N \leq N_{\max}$. We do not apply any additional filtering. We do not exclude values of very low departure throughput from the analysis, because we want to test the hypothesis that low route availability can explain low throughput-high demand observations.

We estimate the departure throughput as a function of arrival throughput and route availability. Given m triplets of measurements $A(t)$, $S_{RAPT}(t)$ and $T(t)$, denoted by $(u_1, w_1, y_1), \ldots, (u_m, w_m, y_m)$, at times when $N^* \leq N \leq N_{\max}$, we seek a function $h_s : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $T = h_s(A, S_{RAPT})$. As before. the constraints are imposed only between neighboring points:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{3.17}$$

subject to:

$$\hat{y}_i = h_s(u_i, w_i), \;\; i = 1, \ldots, m \tag{3.18}$$

$$h_s(i+1, j) \leq h_s(i, j), \;\; i = 0, \ldots (l-1), \forall j, \text{ where } l = \max(A(t)) \tag{3.19}$$

$$h_s(i+1, j) - h_s(i, j) \leq h_s(i, j) - h_p(i-1, j), \;\; i = 1, \ldots (l-1), \forall j \tag{3.20}$$

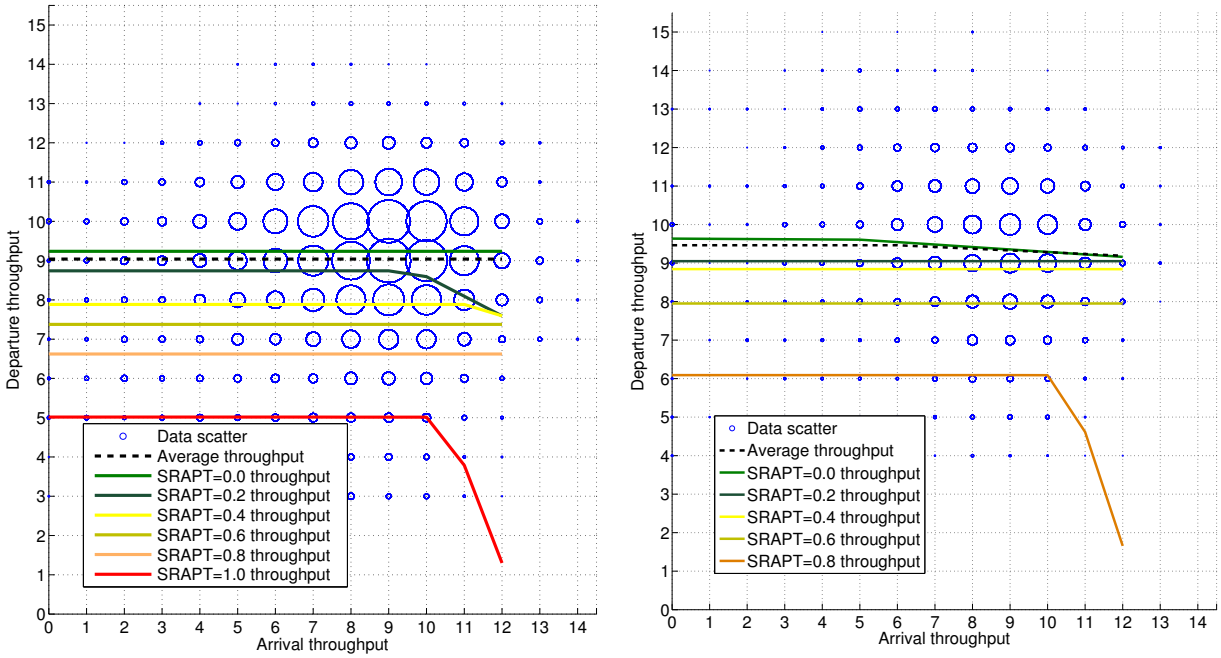
$$h_s(i, j+1) \leq h_s(i, j), \;\; j = 0, \ldots (n-1), \forall i, \text{where } n = \max(S_{RAPT}(t)) \tag{3.21}$$

Inequalities (3.19) and (3.20) are analogous to those in the case of the capacity envelope, i.e., for a given level of route availability, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. Inequality (3.21) ensures that for a given value of arrival throughput, departure throughput decreases as route availability decreases (S_{RAPT} increases). We do not impose more constraints in this fitting problem to avoid making further operational assumptions. For example, it is not clear whether for a given arrival throughput, the departure throughput is a convex, or concave, function of S_{RAPT}.

The estimated function is shown in Figure 3-14. The impact of lower route availability on the departure throughput is evident. We also note that the average throughput is very close to the throughput curve corresponding to the lowest value of S_{RAPT}, that is, clear weather scenarios. It can be seen from Tables 3.15 and 3.16 that the majority of the saturation periods in 2011 were at times with clear routes ($S_{RAPT} = 0.0$).

In conclusion, we also note that "operational throughput", as it was defined, assumes not only persistent demand, but runway availability as well. Therefore, the operational throughput envelope is the one corresponding to the highest route availability ($S_{RAPT} = 0.0$ in this case). If

such information were not available, runway availability would have to be inferred by filtering the data as explained in Section 2.3.1.



(a) Operational throughput envelope for 22 | 13 - configuration, parametrized by the S_{RAPT} value.

(b) Operational throughput envelope for 22 | 31 configuration, parametrized by the S_{RAPT} value.

Figure 3-14: Use of route availability for operational throughput estimation.

3.3.3 Predictive ability of proposed method

We assess the predictive power of the proposed method for the estimation of the impact of the route availability on the capacity of the airport by using five measures for goodness-of-fit: Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAE) and Root Mean Square Percentage Error (RMSPE). We use data from 2011 for training the estimators (Figure 3-14) and we use data from 2010 for testing them. The statistics of the two datasets are listed in Table 3.17.

Table 3.17: Statistics of the training and test datasets for the LGA RAPT analysis.

Runway Configuration	Time frame	Source	Size	Use	Mean	Variance
22 | 13	Jan-Dec 2011	ASPM	26492	Train	8.97	3.04
22 | 13	Jan-Dec 2010	ASPM	31602	Test	8.98	2.92
22 | 31	Jan-Dec 2011	ASPM	8967	Train	9.45	2.86
22 | 31	Jan-Dec 2010	ASPM	11469	Test	9.35	4.04

In order to assess the results, we compare three different estimates for $\hat{y}_i$:

1. The arithmetic mean (AM) of the departure throughput : $y = \frac{1}{m}\sum_{i=1}^{m} y_i$.

2. The throughput predicted from the operational throughput envelope (OTE): $h_s(u_i, 0.0)$.

3. The predicted throughput from the SRAPT-parametrized throughput envelope (RAPT-TE): $h_s(u_i, w_i)$.

We list the statistics for the average operational throughput envelope (OTE), and that parametrized with S_{RAPT} values (RAPT-TE) in Tables 3.18 and 3.19, respectively. We note that RAPT-TE reduces the mean square error in all cases considered. We also note that it provides the greatest variability reduction in the most variable dataset considered, namely, the 22 | 31 throughput in 2010.

Finally, for all goodness-of-fit measures, we report the results for all data for which $S_{RAPT} = 0.0$. The Mean Square Error for this case can be interpreted as the unexplained variation of the operational throughput envelope after filtering-out all time periods with non-clear routes ($S_{RAPT} \neq 0.0$). Comparing it to the MSE of the "OTE", we note that in both training datasets and the testing dataset for configuration 22 | 31, the MSE is reduced 20–40%. This means that route availability explains 20–40% of the variation of the departure throughput of LGA. During time periods with clear routes ($S_{RAPT} = 0.0$), the departure process is much less variable than at other times. Thus, although the point estimates of the departure throughput in saturation (dashed line in Figure 3-14) are not statistically significantly different from the estimates of the departure throughput in saturation given $S_{RAPT} = 0.0$, the variance of the former is much higher.

Table 3.18: Comparison of three different estimators of departure throughput for runway configuration 22 | 13 at LGA, showing the benefits of RAPT usage.

Dataset	Estimator	MAE	MSE	RMSE	MAPE	RMSPE
Training	AM	1.21	3.05	1.75	15.5%	29.4%
Training	OTE	1.21	3.05	1.75	15.5%	29.4%
Training	RAPT-TE	1.29	2.61	1.62	15.01%	26.0%
Training	RAPT-TE, $S_{RAPT} = 0$	1.10	1.92	1.39	12.70%	20.6%
Testing	AM	1.21	2.92	1.71	14.3%	23.4%
Testing	OTE	1.20	2.93	1.71	14.1%	23.3%
Testing	RAPT-TE	1.23	2.77	1.67	14.5%	22.6%
Testing	RAPT-TE, $S_{RAPT} = 0$	1.13	2.31	1.52	13.5%	21.5%

Table 3.19: Comparison of three different estimators of departure throughput on the training data set for runway configuration 22 | 31 at LGA, showing the benefits of RAPT usage.

Dataset	Estimator	MAE	MSE	RMSE	MAPE	RMSPE
Training	AM	1.29	2.86	1.69	16.4%	31.9%
Training	OTE	1.29	2.86	1.69	16.5%	32.3%
Training	RAPT-TE	1.29	2.80	1.67	16.0%	30.5%
Training	RAPT-TE, $S_{RAPT} = 0$	1.27	2.714	1.648	15.5%	29.9%
Testing	AM	1.48	4.04	2.01	18.2%	32.7%
Testing	OTE	1.48	4.14	2.03	18.50%	33.7%
Testing	RAPT-TE	1.42	3.46	1.86	16.78%	27.3%
Testing	RAPT-TE, $S_{RAPT} = 0$	1.29	2.80	1.67	14.3%	21.9%

3.4 Conclusions

In this chapter, we studied the three prominent VMC runway configurations at EWR, and we found that the departure throughput of EWR is not sensitive to changes in runway configuration, arrival throughput and fleet mix.

We then extended the methodology developed in Chapter 2 to study interactions among the three major airports of the NY Metroplex, namely, JFK, EWR and LGA. We found that operations at the three airports are not adversely impacted by operations at the other airports, and we derived capacity envelopes for the system comprising the three airports under different configurations. We estimated that the total balanced operations capacity of the Metroplex is 59 AC/15 min, the departure priority capacity is 53 AC/15 min, and the arrival priority capacity is 63 AC/15 min. We also identified opportunities for performance improvement.

We finally showed that information on route availability can be used for estimating the operational throughput of an airport. We demonstrated that route availability explains a significant fraction of the variation of the departure throughput at LGA, when the prevailing conditions are VMC and the airport is in saturation. In Appendix D, we show that route availability information can be combined with fleet mix information for deriving parametrized operational throughput envelopes for PHL.

Chapter 4

Queuing Model of the Departure Process

In this chapter, we develop an analytical queuing model of the departure process. We train this model using ASPM data from EWR, and evaluate it in terms of its ability to predict taxi-out times and the flow of aircraft on the airport surface. In contrast to Chapters 2 and 3, which focused on estimating the expected throughput under different conditions, this chapter focuses on the derivation of distributions of the random variables involved in the departure process, and the estimation of the impact of their variability on the taxi-out delays.

The main objective of this chapter is to develop a generalizable and easily adaptable model of departure operations. The model development is illustrated for the two main runway configurations of EWR in 2011. The model is also calibrated for PHL in Appendix G, and for CLT in Appendix H.

The model can be used for predicting aggregate taxi-out times and surface congestion, given a pushback schedule for a short, or long time horizon[1]. In this chapter we use the model, which is calibrated using 2011 data from EWR, to predict taxi-out times and surface congestion for departures at the two main runway configurations of EWR in 2007 and 2010. We show that the model can be used for tactical departure planning as well, that is, predicting taxi-out times and departure queues for a short time horizon, like a few hours, or a day. We also assess the impact of different pushback schedules on the variability of delays. Finally, we provide approximate estimates

[1]It is important to note that we do not investigate the impact of uncertainty in the pushback schedules in this work. In other words, we study the predictive properties of the proposed models assuming that the pushback schedules are known. In the current system this may only be realistic for short time horizons (of about 15 minutes).

of the taxi-out times of individual flights.

In addition, the model can be used as a platform for developing and evaluating control algorithms for the departure process, as will be shown in Chapter 5. It is also suitable for policy analysis of infrastructure or operational changes, because of its analytical nature. We describe one such application in Appendix H, where the proposed model is used to assess the impacts of the new runway at CLT on the taxi-out times and the taxi-out delays at the airport.

4.1 Model inputs and outputs

The inputs to the model are

- Pushback schedule, PS.
- Airline of the departing flight, AL.
- Arrival throughput in a 15-minute period starting at time t, $A(t)$.
- Route availability (in the airspace), if available, in a 15-minute period starting at time t, $S_{RAPT}(t)$.
- Segment in use, $(MC; RC)$, expressed as the combination of the visibility conditions, MC, and the runway configuration, RC.

The outputs of the model are

- Number of departures (takeoffs) in the 15-minute period starting at time t, $T(t)$.
- Total number of aircraft taxiing out at the beginning of period t, $N(t)$. It indicates the congestion of taxiing out aircraft on the ground.
- Number of aircraft waiting in the departure queue at the beginning of period t, $Q(t)$. The departure queue is defined as the queue which is formed at the threshold(s) of the departure runway(s), where the aircraft queue for takeoff.
- Number of departing aircraft *traveling* in the ramp and the taxiways towards the departure queue at the beginning of period t (*i.e.*, the number of departures on the surface that have not reached the departure queue), $R(t)$.
- Expected taxi time of departing aircraft l, $\mathbb{E}[\tau(l)]$.

- Expected queuing delay that departing aircraft l experiences, $\mathbb{E}[D_l]$.
- Variance of the queuing delay that departing aircraft l experiences, $var(D_l)$.
- Number of aircraft taking off between the pushback and takeoff time of aircraft l (the length of the takeoff queue experienced by aircraft l [58]), $N_Q(l)$.
- Runway schedule, RW. It refers to the times at which aircraft arrive at the departure queue.

4.2 Model structure

The proposed model, shown in Figure 4-1 consists of two components, or *modules*: (1) the process encompassing aircraft pushing back from the gates and traveling to the departure runway, and (2) the queuing process at the departure runway.

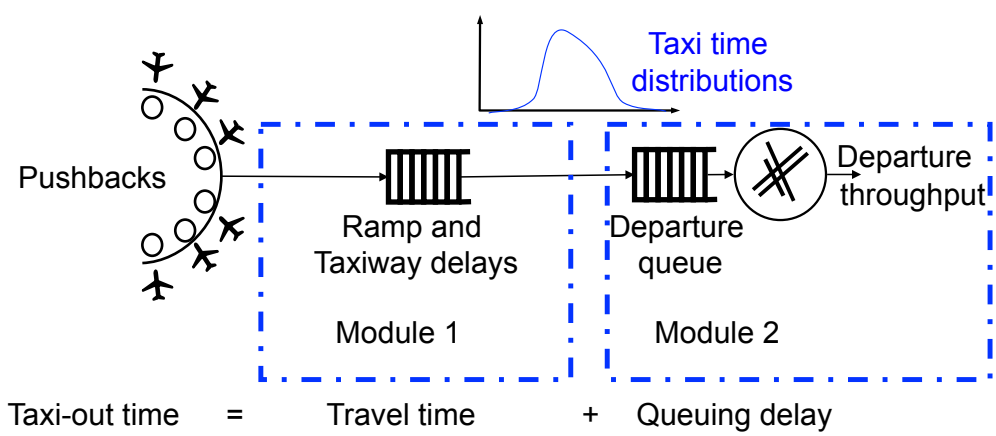


Figure 4-1: Departure process model.

By modeling the departure process in this manner, the taxi-out time $\tau(l)$ of each departing aircraft l can be expressed as

$$\tau(l) = \tau_{travel}(l) + D_l \tag{4.1}$$

where

- $\tau_{travel}(l)$ is the travel time of each departing aircraft l from its gate to the departure runway(s).
- D_l is the queuing delay that aircraft l experiences upon its arrival at the departure queue.

The connection between the two modules is provided by the output of Module 1, that is, the runway schedule.

4.2.1 Data sources

The Aviation System Performance Metrics (ASPM) database offers a wealth of data which enables the study of the performance of the busiest 77 airports in the United States [38]. For every recorded flight, the ASPM database contains their actual push back time and their actual takeoff time. ASPM also reports the runway configuration and the local meteorological conditions at each airport. However, as it has been shown in earlier work [106], the ASPM estimates for the pushback times of the flights that do not automatically report ACARS messages (non-OOOI flights) can be very inaccurate. For these flights, we obtain their pushback times from the Flightstats website [28]. We also obtain information on the terminal and the gate of each flight, from the same website.

4.3 Travel time estimation (Module 1)

In this section, we describe an algorithm that calculates the travel time from the gates to the departure runway(s), namely the first module of the departure process (Figure 4-1). The module can be conceptually described in the following manner: Aircraft pushback from their gates according to the pushback schedule. They enter the ramp and then the taxiway system, and taxi to the departure queue which is formed at the threshold of the departure runway(s). During this traveling phase, aircraft interact with each other. For example, aircraft queue to get access to a confined part of the ramp, to cross an active runway, to enter a taxiway segment in which another aircraft is taxiing, or they get redirected through longer routes to minimize interference with built up congestion. We cumulatively denote these spatially distributed queues and delays, which occur while aircraft traverse the airport surface from their gates towards the departure queue, as *ramp and taxiway interactions*.

We represent the delays due to *ramp and taxiway interactions* with an additive term which gets added to the nominal travel time of each aircraft. The travel time τ_{travel} of each departing aircraft is expressed as:

$$\tau_{travel} = \tau_{unimped} + \tau_{taxiway} \tag{4.2}$$

The first term of Equation (4.2), $\tau_{unimped}$, reflects the nominal or *unimpeded taxi-out time* of the flight. This is the time that the aircraft would spend in the departure process if it were the only aircraft on the ground. The second term, $\tau_{taxiway}$, reflects the delay due to aircraft interactions on the ramp and the taxiways. In other words, $\tau_{taxiway}$ reflects the delay incurred due to other

aircraft that are also on their way to the departure queue. The number of such aircraft is given by $R(t) = N(t) - Q(t)$. The magnitude of this delay will depend on the exact interactions among the taxiing aircraft, that is, the level and location of congestion in the ramp and the taxiways.

4.3.1 Unimpeded taxi-out time

Definition of unimpeded taxi-out times

The unimpeded taxi-out time is the nominal, free flow taxi out time. As the name suggests, it is the taxi-out time of an aircraft if it taxis and takes off in the absence of any obstacles. The FAA defines the unimpeded taxi-out time as *the taxi-out time under optimal operating conditions, when neither congestion, weather nor other factors delay the aircraft during its movement from gate to takeoff* [88]. We note that as per this definition, the unimpeded taxi-out time is not the minimum time that an aircraft would need to taxi-out and take off, but the average time an aircraft needs to complete the departure process when the aircraft spends no time waiting in queues. As explained in earlier work [106], the service time for each of the steps of the departure process is a random variable, and may vary among flights for several reasons, such as:

- Differences during the dispatch stage.
- Routing through different taxiways.
- Different taxi speeds.
- Different runway assignments.
- Variability in the duration of pushback and engine-start.
- Differences in pilot-controller communications.
- Differences in the staffing of the ATCT facility.

There are many forms of delays that an aircraft can incur in this process. For example, there may be communication delays between the pilot and the tower, or the pilot may lack proper weight-and-balance numbers. Factors such as communication delays cannot readily be observed in the recorded data and contribute to the stochasticity of the unimpeded taxi time [18].

Taxi-out times as a function of the adjusted traffic

The unimpeded taxi time is not directly observed, and needs to be estimated. In this section, we propose a new method for the estimation of unimpeded taxi-out times from historical taxi-out times at an airport.

First, we note that unimpeded taxi-out times can vary with the airline, the gate location, the aircraft type[2], the destination[3], the local visibility conditions, the runway configuration, and the runway assignment in case of airports with multiple departure runways [76, 106]. A variable selection for each airport is outside the scope of this chapter. In addition, in all major EWR runway configurations, there is only one departure runway at use. Therefore, for simplicity, we estimate the unimpeded taxi-out time distribution for each airline at given visibility conditions and runway configuration in use at EWR, using ASPM and Flightstats data from 2011. However, the estimation method proposed in this section can be applied to each identified "cluster" of similar flights instead of each airline. An example of such an application is shown in Appendix G for the estimation of unimpeded taxi-out times at PHL.

Based on the findings of Idris et al. [58] and our earlier work [26], we know that the taxi-out time of an aircraft correlates poorly with the number of aircraft taxiing-out on the surface at the time of its pushback. This is because taxiing traffic may impact the aircraft that pushes back differently depending on their relative location, speed and downstream restrictions. Similarly, other aircraft may push *after* the aircraft and yet get ahead of it in the departure queue. For this reason, Idris et al. [58] proposed the concept of the takeoff queue, that is, *the number of aircraft taking off between the pushback and takeoff time of an aircraft*, and developed a probabilistic model to estimate the takeoff queue for each pushback. Subsequently, they used the estimated takeoff queue to predict taxi-out times.

The taxi-out time of an aircraft correlates well with its takeoff queue for several reasons. The larger the number of aircraft that take off while an aircraft is taxiing, the longer that aircraft will have to wait to take off. One may conjecture that this relationship can be used for estimating the unimpeded taxi-out time. When the takeoff queue is zero, there are no aircraft taking off during the time that the aircraft is taxiing, and so it taxies and takes off without incurring any delays due to other aircraft. However, when estimating the takeoff queue from actual data, the estimates are biased: All external factors being equal, an aircraft that ends up with a shorter takeoff queue

[2]Smaller aircraft move at faster speeds.

[3]International flights tend to have longer unimpeded taxi-out times, because of longer checklists [110].

than another will complete the departure process faster than an aircraft that ends up with a longer takeoff queue. For example, if the same aircraft taxies slower, it may get behind another aircraft in the departure sequence. This selection-bias will result in the fastest aircraft having a zero takeoff queue. Thus, an aircraft with a zero takeoff queue will not be representative of *the average aircraft.*

For this reason, we propose the following metrics:

Definition 1 *Departure traffic at time t, $N(t)$, is defined as the sum of the aircraft taxiing out at time t, that is the aircraft that have pushed back, but have not taken off yet.*

Definition 2 *The effective traffic, $N_{eff}(l)$, for an aircraft l is defined as the sum of the aircraft taxiing out, $N(t)$, at the time of its pushback t, and the number of aircraft that push back while it is traveling to the departure runway.*

Based on this, we define the unimpeded taxi-out time as follows:

Definition 3 *The empirical unimpeded taxi-out time distribution of an airline is derived from the values of effective traffic for which the taxi-out time does not increase with increasing effective traffic.*

The effective traffic metric takes into account both the aircraft taxiing-out at the moment of pushback and the traffic that gets added while the aircraft is traveling to the runway. Thus, for low values of $N_{eff}(l)$, it ensures that the aircraft taxies out without being impeded by other aircraft. After the aircraft reaches the queue, it does not matter if additional aircraft push back, as they will not be sequenced in front of it.

The taxi-out time is expected to be a convex non-decreasing function of the effective traffic. Each aircraft needs some amount of time to reach the departure runway. At low values of $N_{eff}(l)$, the existing, or added, traffic has very little probability of interacting with aircraft l. Some aircraft may be behind aircraft l, and some others will takeoff before l reaches the runway. As $N_{eff}(l)$ increases, the probability of other aircraft delaying aircraft l increases. The only aircraft out of $N_{eff}(l)$ that can end up behind aircraft l in sequence, are those that pushed back before aircraft l, but are too slow and the ones that pushed after aircraft l, but did not overtake it. The majority of aircraft in $N_{eff}(l)$ will be the *takeoff queue* for aircraft l. As $N_{eff}(l)$ increases even further, each additional aircraft will impose a delay approximately equal to its service time at the runway. Thus the slope of the function will increase from 0 to the average service time at the runway.

We note here that the definition of effective traffic assumes availability of ASDE-X data. ASPM data do not provide information on the times that aircraft join the departure queue. For ASPM data, we heuristically modify the $N_{eff}(l)$ metric with the following one, first proposed by Clewlow [95]:

Definition 4 *The adjusted traffic, $N_{adj}(l)$, for an aircraft l is defined as the sum of the aircraft taxiing out, $N(t)$, at the time of its pushback t, and the number of aircraft that push back while aircraft l is taxiing out.*

The two definitions ($N_{eff}(l)$ and $N_{adj}(l)$) are practically equivalent for estimating the taxi-out time at low values of traffic.

Figure 4-2 shows the relationship between the adjusted traffic and the taxi-out time for flights of ExpressJet Airlines[4]. We show the scatter plot, along with the mean, standard deviation and the median taxi-out times for all values of the adjusted traffic. We notice the non-linear relationship between the adjusted traffic and taxi-out time. Despite relaxing the definition of the effective traffic to include also aircraft that push while aircraft l is in the departure queue, the mean taxi-out time and the median taxi-out time appear to remain convex non-decreasing functions of the adjusted traffic. This relation was observed in all cases considered. We therefore conjecture that taxi-out times are a convex function of the adjusted traffic.

Finally, in Figure 4-2, we note that the median values are consistently lower than the mean values. Similarly, we observe that the data-scatter for the each value of the adjusted traffic has a longer tail for higher values of taxi-times. This is to be expected as taxi-out times, even the unimpeded ones, can only be as short as the physics of the process allows, but they can grow large, depending on a slow pushback. This is also consistent with the literature that suggests Erlang, or Lognormal distributions for the unimpeded taxi-out time [20, 70].

For fitting a curve to the observations of Figure 4-2, we propose a simple estimation program. We want to fit a convex, non-decreasing function to the data that estimates the taxi-out time as a function of the adjusted traffic. Given m pairs of measurements $N_{adj}(l)$ and $\tau(l)$, denoted $(u_1, y_1), \ldots, (u_m, y_m)$, we seek a convex non-decreasing function $f : \mathbb{R} \to \mathbb{R}$ that estimates the mean $\tau = f(N_{adj}(l))$. This infinite-dimensional problem is significantly simplified by the fact that N_{adj} is defined only in the domain of natural numbers ($\mathbb{N}_0$). f can be restricted to within the domain of $\mathbb{N}_0$ as well, and we need to estimate the values $f(0), f(1), \ldots, f(n)$, where $n = \max(N_{adj})$. The function

[4]regional partner of Continental/ United Airlines

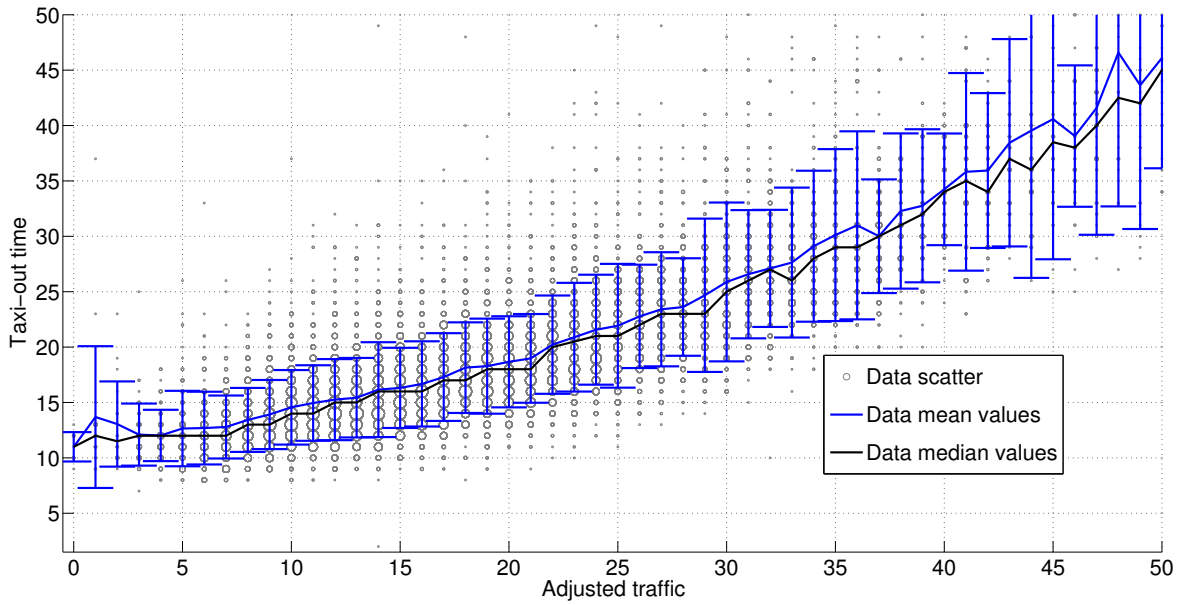


Figure 4-2: Empirical data showing the taxi-out times as a function of the adjusted traffic for the flights of ExpressJet Airlines in configuration 22L | 22R at EWR.

f is simply a piecewise linear function of N, and the monotonicity and convexity constraints are imposed at the points $0, 1, \ldots, \max(N_{adj})$ by comparing the values and the slopes of subsequent pieces. f is given by the solution to the following convex optimization problem:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{4.3}$$

subject to:

$$\hat{y}_i = f(u_i), \;\; i = 1, \ldots, m \tag{4.4}$$

$$f(i+1) \geq f(i), \;\; i = 0, \ldots (n-1) \tag{4.5}$$

$$f(i+1) - f(i) \geq f(i) - f(i-1), \;\; i = 1, \ldots (n-1) \tag{4.6}$$

The results of the regression fit are shown in Figure 4-3. We notice that the estimated function simply smooths out the mean values of the raw data. We also note here that we can impose constraints on the number of breakpoints and force the fitted function to comprise a certain number of linear segments by applying piecewise regression. However, in this case, we do not want to add more constraints on the set of values of N_{adj} for which the taxi-out time does not increase with increasing N_{adj}. By inspecting Figure 4-3, we recognize that the fitted function is flat for $0 \leq N_{adj} \leq 4$. Thus, the observations of taxi-out times for which $0 \leq N_{adj} \leq 4$ provide the

empirical distribution of the unimpeded taxi-out times for ExpressJet Airlines. We fit a log-normal distribution to the empirical distribution with mean equaling the estimated value in the flat region ($f(0)$) and standard deviation equaling the standard error of the estimated mean. The resulting distribution and the empirical one are shown in Figure 4-4.

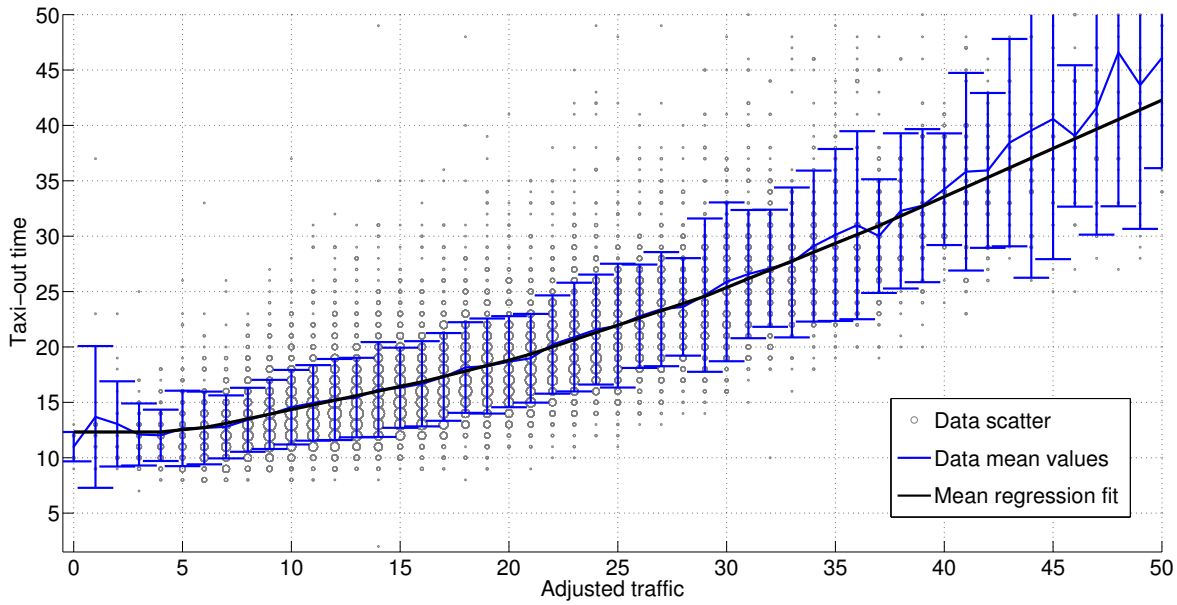


Figure 4-3: Regression fit of the taxi-out times as a function of the adjusted traffic for the flights of ExpressJet Airlines in runway configuration 22L | 22R at EWR.

We show the estimated functions for two more airlines, JetBlue and US Airways for this runway configuration of EWR in Appendix E. We also show a table with the parameters of the unimpeded taxi-out time distributions for all major carriers of EWR in runway configuration 22L | 22R. In Section 4.5.2, the distributions of the predicted taxi-out times are compared to the actual taxi-out times. From comparing the distributions at low values of traffic, we validate that the unimpeded taxi-out times are estimated correctly.

4.3.2 Ramp and taxiway interactions

In earlier work [107], we proposed a simple formula for modeling ramp and taxiway interactions:

$$\tau_{taxiway} = \alpha \cdot R(t) \tag{4.7}$$

Equation (4.7) implies:

$$\tau_{travel} = \tau_{unimped} + \alpha \cdot R(t) \tag{4.8}$$

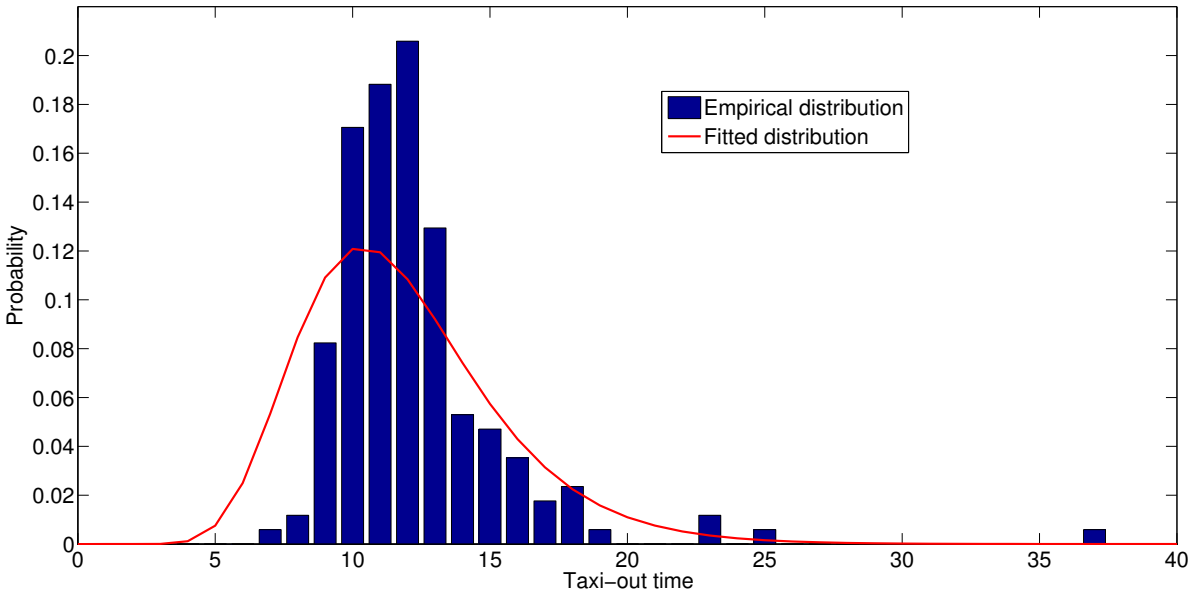


Figure 4-4: Empirical and fitted distribution of the unimpeded taxi-out times of ExpressJet Airlines in runway configuration 22L | 22R at EWR.

In Equations (4.7) and (4.8), the term $\alpha R(t)$ is a linear term used to model the interactions among departing aircraft on the ramps and taxiways. α is a parameter that depends on the airport and the runway configuration. The presence of this term is justified, since aircraft are expected to stop at intersections on the ramp and taxiways. The probability of stopping increases with the number of aircraft traveling towards the departure queue, at time t of pushback, $R(t)$. The parameter α can be interpreted as the expected number of stops multiplied by the expected length of each stop. For instance, consider the expected number of stops being 0.2 multiplied by the number of aircraft traveling. Thus if 10 aircraft are traveling at the time aircraft l pushes back, aircraft l will experience on average 2 stops. If a stop lasts 0.5 min, this would translate to an expected delay of 1 min. In this example, the parameter α equals $0.5\,\text{min} \times 0.2/\text{AC} = 0.1\,\text{min}\,/\text{AC}$. In earlier work, Equation (4.8) was validated using ASPM data and was subsequently successfully applied in several studies regarding a variety of airports layouts [76, 85, 113]. We note that it is a very aggregate way to describe interactions among aircraft in the ramp and queue areas, as different areas of the tarmac imply higher probabilities of stopping (for example the horseshoe ramp areas in EWR, BOS, PHL airports and others). Nonetheless, Khadlikar validated the linear dependence of the travel time with the level of traffic using ASDE-X data [70], by studying the dependence of the time an aircraft spends traversing a certain link of the airport surface with the level of traffic

on the surface. Using ASDE-X data, we have also shown that by neglecting ramp and taxiway interactions, the rate at which aircraft arrive at the runway is overestimated [110].

In this work, we do not use ASDE-X data. We select parameter α using a data-based method. We choose α such that we match the predicted median taxi-out time with the actual median taxi-out time. The reason for this choice is that we do not have information on traffic management initiatives which tend to significantly delay aircraft for reasons not directly related to the departure process of the origin airport [59], so we expect the estimated mean taxi-out time from our model to be shorter than the actual mean taxi-out time. However, the median taxi-out time is less sensitive to these outliers with very long taxi-out times.

4.3.3 Module 1 output: Expected runway schedule

In Section 4.2, we saw that Module 2 uses as input the runway schedule, which is provided by Module 1. In reality, the runway schedule is not known in advance, as the travel times from the gates to the departure queue are random variables. In the model, we assume that aircraft have fixed unimpeded taxi-out times, which are equal to their expected values. In other words, we assume that aircraft progress to the departure queue with their expected travel times.

$$\mathbb{E}[\tau_{travel}(l)] = \mathbb{E}[\tau_{unimped}(l)] + \alpha \cdot R(t_{pb}(l)) \tag{4.9}$$

In Equation (4.9), $\mathbb{E}[\tau_{unimped}(l)]$ is the mean unimpeded taxi-out time of the airline that operates aircraft l and $R(t_{pb}(l))$ is the number of traveling aircraft at the time of pushback, $t_{pb}(l)$, of aircraft l. We note that $R(t)$ is calculated as the program progresses. Each aircraft increases R by one unit during the time it progresses to the runway. In this way, the *known* runway schedule is generated, and this is the input to the runway(s) in Module 2. Module 2 calculates the expected queuing delays given this runway schedule. From the combined output of the model we get:

$$\mathbb{E}[\tau(l)] = \mathbb{E}[\tau_{unimped}(l)] + \alpha \cdot R(t_{pb}(l)) + \mathbb{E}[D_l] \tag{4.10}$$

The rationale for making the simplifying assumption that aircraft have fixed unimpeded taxi-out times is that their unimpeded taxi-out time distributions are in general concentrated around their expected value (Table E.1). Our hypothesis is that, provided the pushback times are known, the variability of the unimpeded taxi-out times introduces very little additional delay. We verify, through simulation, that the hypothesis of a *known* runway schedule for calculating delays is indeed

a good approximation (Section 4.5.2).

4.4 Queuing delay estimation (Module 2)

In this section, we discuss the queuing model used for predicting the delays at the departure runway(s). In Chapters 2 and 3, it was shown that the departure throughput of an airport is a dynamic and stochastic process. External conditions (arrival throughput, downstream restrictions, etc.) dynamically change the operational characteristics of the departure process, but do not explain all of its variability. As was seen in Table 3.15, there is still unexplained variability in the departure throughput at LGA after controlling for the arrival throughput and the route availability. The departure process is a probabilistic process, the unpredictable variability of which stems from a number of factors, such as controllers' and pilots' decisions, aircraft performance, human errors, and incidents like aborted takeoffs, runway closures, etc.

4.4.1 Runway queuing model

We propose an analytical queuing model for estimating the queuing delays during the departure service process of an aircraft. The proposed approach can be used to model the whole runway system, or each individual runway, depending on the data availability, and the desired level of modeling.

We define as the service rate, $\mu(t)$, the number of departing aircraft that can take off from the runway(s) modeled per 15-minute interval. Four fundamental assumptions are made in our attempt to model the service process in the runway system:

1. The demand is given by the known runway schedule, calculated by Module 1 of the model. Thus, aircraft arrive at the queuing system according to the known runway schedule.
2. The service rate is assumed to follow a time-dependent (dynamic) Erlang distribution.
3. There is finite queuing space for the departing aircraft to wait in.
4. Aircraft in the departure queue are served on a First-Come-First-Served (FCFS) basis.

This framework is attractive because it models both the mean and variance of the departure process throughput, which have been shown to be important when calculating queuing delays in the context of airport systems [53, 54]. In queuing theory, this type of system would be denoted as an $D(t)/E_k(t)/1$ with finite queuing space C. In our case, as it will be explained later, the finite queuing space C is introduced for computational tractability and not for modeling the finite

available space for aircraft to queue around the departure runway(s).

In terms of the queuing literature for airport systems modeling, this model fits into the "Micro Model" literature. Each period of time of a day that the runway configuration considered is in use is divided into time windows of equal duration, Δ, each of 15 minutes, and indexed with $i = 1, 2, \ldots, T$. For each time period i, a throughput distribution, of type Erlang with rate $k_i\mu_i$ and shape k_i, is provided based on the operating conditions in the airport. Each flight of the runway schedule is indexed with index l. During each time window i, we have (from Module 1) the set of arrival times at the departure queue of aircraft in this time window, $S(i)$. Their arrival times are provided by Module 1 and are assumed known a priori, but are not necessarily uniformly spaced in each time window.

We also note here that Module 2 models one departure runway in the case of EWR (Runway 4L/22R). Thus, the modeled queuing system maps to the physical queuing process at Runway 4L/22R. If there are multiple runways and the runway assignments are known, we can model each of them with a queuing system. An application of such a model for the departure process at Detroit Metropolitan Wayne County Airport (DTW) can be found in other work [76]. By contrast, in Appendix G, we use a single queuing system to model the queuing process at the main runway (27L) and the secondary runway (35) at PHL. Similarly in Appendix H, we use a single queuing system for modeling the departure runways 18C/36C and 18L/36R at CLT. The approximation of modeling two departure runways with a single queuing system is driven by the lack of runway assignments in the ASPM data. However, the predictive power of the model remains strong, as can be verified in Appendix G and in Appendix H.

We model this system with a discrete-time Markov Chain. A service completion of an Erlang process with shape k and rate $k\mu$ is represented with k stages of exponentially distributed random variables with rate $k\mu$. We call each such stage a *stage-of-work*. Each stage of the Markov chain q denotes that there are q *stages-of-work* to be completed at the runway, i.e., there are $\min(1, q)$ aircraft in service and $\max(\lceil (q-k)/k \rceil, 0)$ aircraft in the departure queue.

We summarize the notation used in this section:

- l: Index of each aircraft.
- Δ: Duration of each time window.
- i: Index of each time window.
- k_i: Shape parameter of the Erlang service time distribution during time window i.

- $k_i\mu_i$: Rate parameter of the Erlang service time distribution during time window i.
- C: Queuing space of the queuing system, measured in units of aircraft.
- Q_l : *stages-of-work* to be completed for emptying the system immediately after the lth aircraft's arrival. This means that there are Q_l stages-of-work to be completed (including its service time) until the lth aircraft exits the system.
- $p_{Q_l}(j)$: Probability that Q_l takes the value j.
- c_l : Inter-arrival time at the system between the $(l-1)^{\text{th}}$ and l^{th} aircraft.
- $C_l = \sum_{j=1}^{l} c_j$: time of arrival of the lth aircraft at the system.
- S_l: Total time that aircraft l spends in the queuing system, including time-in-service.
- D_l: Queuing delay that aircraft l experiences.
- Z_0: Number of aircraft in queue at the beginning of the first time window.
- $f_\nu(x)$: P.m.f. of random variable X, drawn from a Poisson distribution with parameter ν.
- $F_\nu(x)$: C.d.f. of random variable X, drawn from a Poisson distribution with parameter ν.

Static service process distribution

Here, we assume that that service process is static and is described by a single Erlang distribution with parameters $(k, k\mu)$ at all time windows. We observe the system at the epochs of arrival, C_l, of each aircraft l. In Figure 4-5, we show an example of the state of the Markov chain that an aircraft encounters, assuming that there are exactly 3 *stages-of-work* to be completed before the arrival of lth aircraft in the system, and an Erlang distribution with shape 2 for the departure throughput. Upon the arrival of lth aircraft, 2 stages-of-work are added and the system (instantaneously) transitions to state 5.

We turn now to calculating the probabilities $p_{Q_l}(j)$. p_{ij}^l is defined as the probability of the following event: There are i stages-of-work to be completed immediately after the arrival of the lth aircraft, given that there were j stages-of-work to be completed immediately after the arrival of $(l-1)^{\text{th}}$ aircraft:

$$p_{ij}^l = \Pr(Q_l = i | Q_{l-1} = j) \tag{4.11}$$

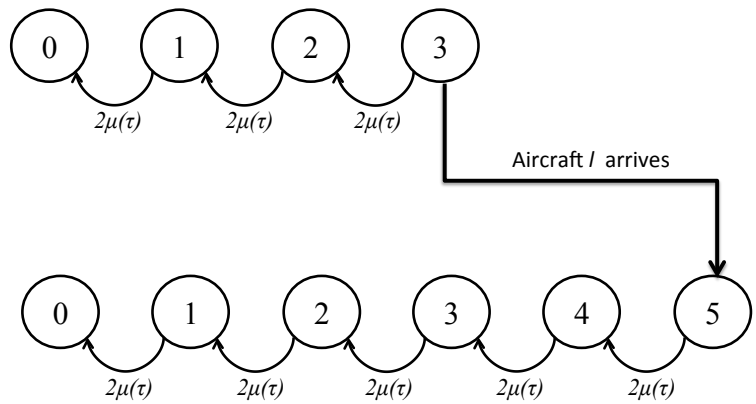


Figure 4-5: Markov chain transition at the time of arrival of the lth aircraft.

If the system is at state Q_{l-1} when $(l-1)^{\text{th}}$aircraft arrives, it evolves during time c_l similarly to a Poisson process:

$$p_{ij}^{l} = \begin{cases} \frac{e^{-k\mu c_l} \cdot (k\mu c_l)^{j+k-i}}{(j+k-i)!} & \text{if } i = k+1, k+2, \ldots, j+k, i < kC, i \leq j-k \\ 1 - \sum_{z=k+1}^{j+k} \frac{e^{-k\mu c_l} \cdot (k\mu c_l)^{z-k}}{(z-k)!} & \text{if } i = k \\ \sum_{z=0}^{j-i+k} \frac{e^{-k\mu c_l} \cdot (k\mu c_l)^{z}}{(z)!} & \text{if } i = kC, i \leq j-k \\ 0 & \text{otherwise} \end{cases} \tag{4.12}$$

Equation (4.12) can be explained as follows: Define as an *event*, the completion of a stage-of-work in the Markov chain of the system. The time between each pair of consecutive events has an exponential distribution with parameter $(k\mu)$. Each of the inter-arrival times of *events* is identically distributed and independent of other inter-arrival times. Thus, the probability of the completion of $i < j$ stages-of-work (out of j) in time c_l is given by the Poisson distribution with parameter $(k\mu c_l)$. We note that up to j stages-of-work can be completed, and thus we can have up to j *events* during time c_l. The condition $i \leq j$ does not impact the probability distribution of $i < j$ *events*. Each event is independent of its previous and its future ones. The condition $i \leq j$ impacts the probability distribution of having exactly $i = j$ *events* (emptying the queuing system), which is given by the probability of having $j, j+1, j+2, \ldots$ Poisson arrivals (or one minus the probability of having $1, 2, \ldots, j-1$ *events*).

For the third case of Equation (4.12) we have: If the system is in a state higher than $kC - k$ after the arrival of the $(l-1)^{\text{th}}$ aircraft, we enforce that there will be space for the lth aircraft when it arrives. Thus, for $k(C-1) < i \leq kC$, the system transitions to state j with the sum

of the probability of $0, 1, \ldots, i-j+k$ *events*. In other words, we enforce the completion of as many stages-of-work as necessary to make space for the arrival of lth aircraft. We note that this is different from standard queuing systems modeling convention, where the extra customer is rejected from the system if there is no queuing space. Our convention gives an opportunity for biasing the performance of the system, since selecting a very low value of parameter C can "boost" the throughput of the system, by forcing stage-of-work completions at each arrival epoch. Thus, selecting the queuing space parameter C is governed by tradeoffs between computational tractability and accuracy. Numerical experiments suggest that setting $C = 100$ is a good choice for modeling the runway queuing system. In other words, for $C = 100$, the system behaves as if it had infinite queuing space.

During time c_l, the system cannot get to a state higher than i since there are no arrivals of aircraft in the system. We also note, that the calculated probabilities of having $0, 1, \ldots, j$ *events* during time c_l map to the probabilities $p^l_{(j+k)j}, p^l_{(j+k-1)j}, \ldots, p^l_{kj}$ because the arrival of lth aircraft at time C_l instantaneously adds k stages-of-work.

Using the Poisson distribution notation introduced, Equation (4.12) can be written in the following form, where $\nu = k\mu c_l$:

$$P(\nu) = \begin{array}{c} \\ 0 \\ \vdots \\ k-1 \\ k \\ k+1 \\ \vdots \\ kC-1 \\ kC \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & \cdots & kC-k & kC-k+1 & \cdots & kC-1 & kC \end{array} \\ \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1-F_\nu(0) & \cdots & 1-F_\nu(kC-k-1) & 1-F_\nu(kC-k) & \cdots & 1-F_\nu(kC-2) & 1-F_\nu(kC-1) \\ 0 & f_\nu(0) & \cdots & f_\nu(kC-k-1) & f_\nu(kC-k) & \cdots & f_\nu(kC-2) & f_\nu(kC-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & f_\nu(1) & f_\nu(2) & \cdots & f_\nu(k) & f_\nu(k+1) \\ 0 & 0 & \cdots & F_\nu(0) & F_\nu(1) & \cdots & F_\nu(k-1) & F_\nu(k) \end{pmatrix} \end{array} \tag{4.13}$$

Using Equation (4.12) we calculate the probabilities p_{Q_l}:

$$p_{Q_l}(i) = \sum_{j=(i-k)}^{j=kC} p^l_{ij} p_{Q_{l-1}}(j) \text{ for } i = k, k+1, \ldots kC \tag{4.14}$$

$$\text{or } \mathbf{p}_{Q_l} = P(\nu)\mathbf{p}_{Q_{l-1}} \tag{4.15}$$

where $\mathbf{p}_{Q_l} = [p_{Q_l}(0), p_{Q_l}(1), \ldots, p_{Q_l}(kC)]'$

The starting condition for the calculation is the queue that the first aircraft arrival encounters

(Z_0). If the queue is empty at the beginning of the time period modeled, $Z_0 = 0$, and the first aircraft will bring the system to state k upon its arrival. A non-empty queue starting condition is interesting for runway configuration changes, where aircraft that pushed back before the runway configuration change are lined up at the new departure runway when its use commences.

$$p_{Q_0}(i) = \begin{cases} 1 & \text{if } i = k \cdot Z_0 \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

By inspection, we note that Matrix (4.13) is a stochastic matrix, as its columns add up to 1 for all values of k, and ν. Thus for a valid $\mathbf{p}_{Q_{l-1}}$ vector, the multiplication of Equation (4.15) returns a valid probability vector $\mathbf{p}_{Q_l}$.

To summarize, Equation (4.16) gives the state of the queue at the first aircraft's arrival at time C_1. The state of the queue at the times of arrival of aircraft $2, 3, \ldots$ is calculated with Equation (4.15).

Dynamic service process distribution

For dynamic service time distributions, Equation (4.15) is still valid by using the appropriate service time distribution parameters $(k_i, k_i\mu_i)$ for the time window i, in which inter-arrival time c_l falls:

$$\nu_i = k_i \mu_i c_l \tag{4.17}$$

$$\mathbf{p}_{Q_l} = P(\nu_i)\mathbf{p}_{Q_{l-1}} \tag{4.18}$$

A subtlety here is that the state space of the queuing system can change. For example, if we have two throughput distributions, with shapes k_1 and k_2, where $k_1 \neq k_2$, the space of the chain of the queuing system corresponding to the first distribution will be $\{0, 1, \ldots, k_1 \cdot C\}$ and the second $\{0, 1, \ldots, k_2 \cdot C\}$. Transitioning from the throughput distribution of the former to that of the latter can be done only by approximately mapping the probabilities of states $\{0, 1, \ldots, k_1 \cdot C\}$ to those of $\{0, 1, \ldots, k_2 \cdot C\}$.

4.4.2 Queuing delay calculations

Static service process distribution

Given the probabilistic state of the Markov chain described by $\mathbf{p}_{Q_l}$ at the time of arrival, we calculate the moments of the queuing time of each aircraft, assuming a static service process, that is, an Erlang distribution with parameters $(k, k\mu)$.

$$\begin{aligned}
\mathbb{E}[S_l] &= E\big[\mathbb{E}[S_l|Q_l]\big] && (4.19)\\
&= E\big[\frac{j}{k\mu}|Q_l = j\big] && (4.20)\\
&= \sum_{j=0}^{kC} \frac{j \cdot p_{Q_l}(j)}{k\mu} && (4.21)
\end{aligned}$$

Similarly:

$$\begin{aligned}
\mathbb{E}[S_l^2] &= E\big[\mathbb{E}[S_l^2|Q_l]\big] && (4.22)\\
&= E\big[\frac{j(j+1)}{\mu^2}|Q_l = j\big] && (4.23)\\
&= \sum_{j=k}^{kC} \frac{j(j+1)}{(k\mu)^2} \cdot p_{Q_l}(j) && (4.24)
\end{aligned}$$

In Equations (4.21) and (4.24), we note that given that the system is in state j immediately after the arrival of the lth aircraft, the total time that the lth aircraft spends in the system is given by the sum of j exponential distributions, each with rate $k\mu$. The probability distribution for its delay is given by an Erlang distribution with shape j and rate $k\mu$. Thus, given the state j at time of arrival, we can fully characterize the time-in-the-system distribution for aircraft l. If the state j is not known, we can derive all the moments of the distribution in the system using the total expectation theorem and the moment generating function for the Erlang distribution, as shown in Equations (4.21) and (4.24) for the first two moments. This process can be computationally expensive, so we restrict our efforts to the first two moments, which yield the expected delay and its variance for each aircraft. We also note that Equations (4.21) and (4.24) refer to the total time that aircraft l spends in the queuing system. For calculating the queuing delay we need to subtract

the time that is spent for servicing l^{th} aircraft, that is, the last k stages-of-work.

$$\begin{aligned}
\mathbb{E}[D_l] &= E\big[\mathbb{E}[D_l|Q_l]\big] && (4.25)\\
&= E\big[\frac{j-k}{k\mu}|Q_l = j\big] && (4.26)\\
&= \sum_{j=k}^{kC} \frac{(j-k)\cdot p_{Q_l}(j)}{k\mu} && (4.27)\\
\mathbb{E}[D_l^2] &= E\big[\mathbb{E}[D_l^2|Q_l]\big] && (4.28)\\
&= E\big[\frac{(j-k)(j-k+1)}{(k\mu)^2}|Q_l = j\big] && (4.29)\\
&= \sum_{j=0}^{kC} \frac{(j-k)(j-k+1)}{(k\mu)^2}\cdot p_{Q_l}(j) && (4.30)
\end{aligned}$$

Finally, the probability vector $\mathbf{p}_{Q_l}$ can be used to calculate the probability, $r_l(\bar{d})$, of the queuing delay of flight l exceeding a certain threshold $\bar{d}$: [5]

$$\begin{aligned}
r_l(\bar{d}) = \Pr(D_l > \bar{d}) &= \sum_{j=k}^{kC} \Pr(D_l > \bar{d}|Q_l = j)\cdot p_{Q_l}(j) && (4.31)\\
&= \sum_{j=k}^{kC} \Big(1 - \frac{\gamma(j, k\mu\bar{d})}{(k\mu-1)!}\Big)\cdot p_{Q_l}(j) && (4.32)
\end{aligned}$$

Dynamic service process distribution

In the general case, the service process is dynamic. This means that parameters $(k, k\mu)$ can change across time windows. In this case, the Equations (4.27), (4.30) and (4.32) cannot be used, because the intermediate Equations (4.26), (4.29) and (4.31) do not apply anymore. Given a state in the queuing system we cannot calculate the expected delay, because the parameters of the associated transition change dynamically. This problem has been acknowledged by other authors working with dynamic models [8, 53, 93]. Here, we propose the following approximation, building on Gupta's approximation for D(t)/M(t)/1 systems.

For each aircraft, we calculate the *effective queue*, $\tilde{q}_i(j)$, that aircraft l encounters immediately after entering the queuing system:

$$\tilde{q}_l(j) = \sum_{j=k_i}^{k_iC} \frac{(j-k_i)\cdot p_{Q_l}(j)}{k_i} \qquad (4.33)$$

[5] This metric for evaluating the risk of high delays was recently proposed by Jacquillat [63].

The effective queue, $\tilde{q}_l(j)$, is an estimate of the *physical queue* that aircraft l faces. Equation (4.33) measures the queue in terms of aircraft and not *stages-of-work*. It does not depend on the service time distributions in the future. In addition, it is a more accurate estimate of the queue than the *expected queue*, $\mathbb{E}[Q_l]$:

$$\mathbb{E}[Q_l] \quad = \quad \sum_{j=k_i}^{k_i C} \lceil \frac{j-k_i}{k_i} \rceil \cdot p_{Q_l}(j) \tag{4.34}$$

This is because for the expected queue calculation, all *stages-of-work* associated with one unit in the *physical queue* collapse to a single state. By contrast, when calculating the *effective queue*, each stage-of-work is considered separately.

For each aircraft arriving at the departure queue, we calculate its effective queue, $\tilde{q}_l(j)$, at the moment of its arrival. Subsequently, we use the equivalent deterministic service process to calculate the total time that it will take for the effective queue, $\tilde{q}_l(j)$, to be dissipated. In this way, each aircraft in the effective queue of aircraft l is served with the service rate that applies at the time of their takeoff. We call this time *effective delay*, and denote it $\tilde{d}_l(j)$. To demonstrate the effective delay idea, we give two examples, assuming that each time window has a duration of 15 minutes.

In the first example, aircraft 1 arrives at the queuing system in the 5th minute of Δ_1. During Δ_1 the service rate is 10 AC/15 min. We calculate the effective queue for aircraft 1 and it is found to be 4 AC. Given the current service rate, all of the 4 AC will be served in time window Δ_1 assuming a deterministic service process. It will take 6 minutes (4AC / (10 AC/15 min)) to serve them and this is the adjusted queuing delay.

In the second example, aircraft 1 arrives at the queuing system in the 5th minute of Δ_1. In Δ_1, the service rate is 10 AC/15 min, and in Δ_2, the service rate is 8 AC/15 min. We calculate the effective queue for aircraft 1 , which equals 10 AC. Assuming a deterministic service process, 6.7 of them are served in the remainder of Δ_1 (10/15 $\times$10 min). The remaining 3.3 AC are served in time window Δ_2 with a service rate of 8 AC/15 min, which corresponds to a delay of 6.25 min. Thus the total effective delay equals 10 min + 6.25 min, or 16.35 min.

We note that the notion of the effective delay follows very closely Gupta's proposition for the *effective service rate* [53]. The contribution of our approach is the notion of the *effective queue*, which is necessary for considering Erlang service time distributions.

For estimating the variance of the queuing delays, we make yet another approximation and simply use Equation 4.30, with parameters $(k, k\mu)$ those at the time of arrival of the aircraft at the

departure queue.

4.4.3 Estimation of the departure capacity distributions

For estimating the departure process characteristics we follow an approach similar to that described in Chapters 2 and 3 for estimating the operational throughput envelopes.

For EWR's most frequently used runway configuration, (VMC; 22L | 22R) in 2011, we first plot the saturation plot, shown in Figure 4-6. The saturation plot shows the dependence of the departure throughput on the departure demand. We note that the average throughput stabilizes at around 17 aircraft taxiing out and stays at around 10 AC/15 min until it starts fluctuating for taxiing out traffic higher than 25 aircraft on the ground. We also note that the departure throughput has very high variability. For measuring the impact of other explanatory variables, we use the approach outlined in Section 2.3.1: We construct the regression tree that shows the departure throughput (T) as a function of the departure demand (N), the arrival throughput (A) , and the surrogate RAPT value (S_{RAPT}). The regression tree indicates that EWR is in saturation for $16 \leq N \leq 31$.

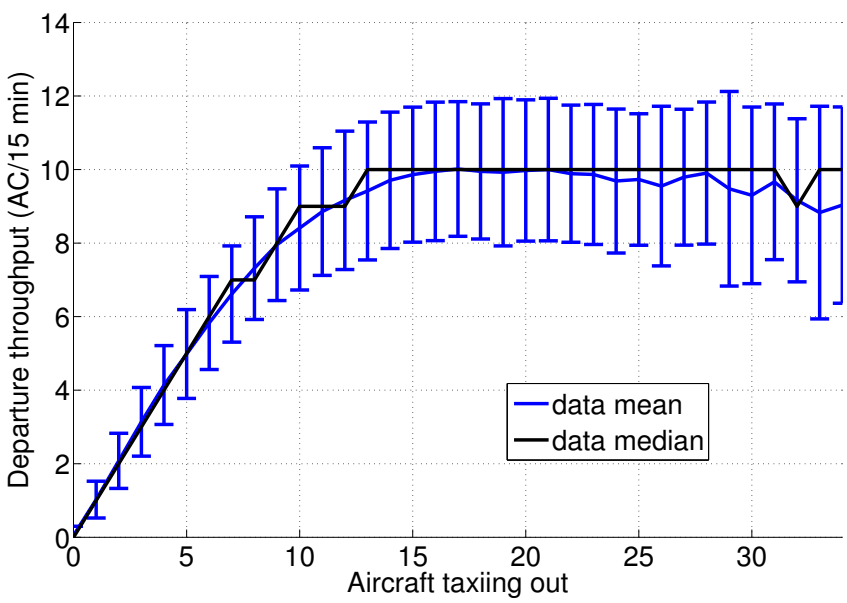
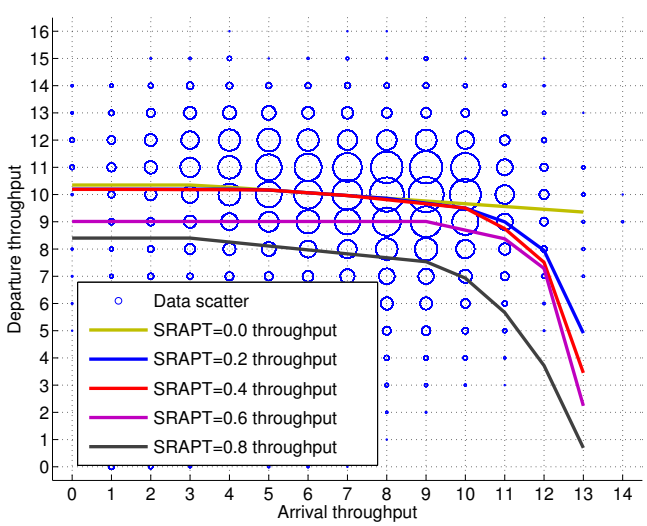


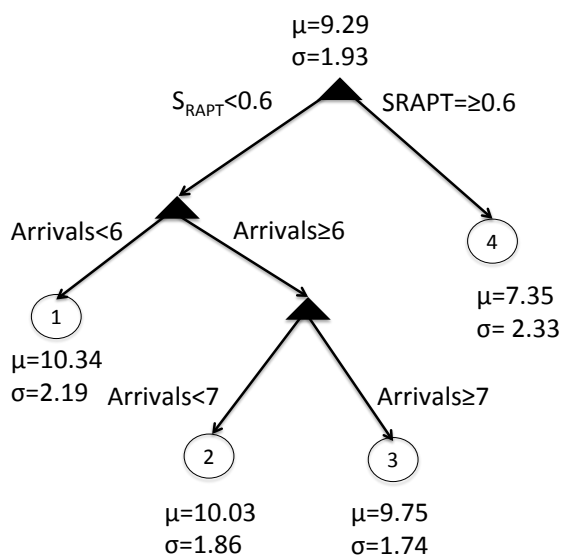
Figure 4-6: EWR saturation plot for configuration (VMC; 22L | 22R) in year 2011.

We use the saturation data-points to estimate the departure throughput as a function of the arrival throughput and the route availability (RAPT). We first use the methodology outlined in Section 3.3.1 for estimating the departure throughput as a function of route availability and arrival throughput. In Figure 4-7a, we show the departure throughput curves parametrized by route availability. Each point of these curves provides a point estimate for the expected departure throughput

given an arrival throughput and a S_{RAPT} value. These point estimates are very useful for representation purposes, but do not provide information on higher order moments of the throughput distribution. For this reason, we turn to regression tress. We estimate the departure throughput in saturation as a function of arrival throughput and route availability. The resulting (pruned) tree is shown in Figure 4-7b. Comparing the throughput curves with the regression tree, we note that in the regression tree, several operational points of the throughput curves are merged in a single leaf of the tree. For example, all points of the curves with S_{RAPT} values lower than 0.6 for which arrival throughput is less than 6 AC/15 min are represented by one leaf with throughput 10.34. From their curves in Figure 4-7a, we note that for arrival throughput less than 6 AC/15 min and S_{RAPT} value lower than 0.6, all throughput estimates are between 10 AC/15 min and 10.5 AC/15 min. Given the variance of the departure throughput measurements in this area, all these estimates are merged into a single leaf with expected throughput (μ) 10.34 AC/15 min and standard deviation (σ) 2.19 AC/15 min.



(a) Throughput curves for (VMC; 22L | 22R) configuration parametrized by route availability.



(b) Regression tree showing the departure throughput of (VMC; 22L | 22R) configuration parametrized by route availability and arrival throughput.

Figure 4-7: Two representations of the EWR departure capacity

Erlang distribution fitting

We use trees such as the one of Figure 4-7b for characterizing the dynamics and stochasticity of the departure process. At each 15-minute period, the route availability and the arrival throughput

map to a leaf of the tree of Figure 4-7b. Each leaf of the tree is associated with an empirical distribution that consists of all the observations satisfying its conditions. These distributions are denoted f_{rw}. Figure 4-8 shows the corresponding four empirical distributions f_{rw} for each leaf of the tree in Figure 4-7.

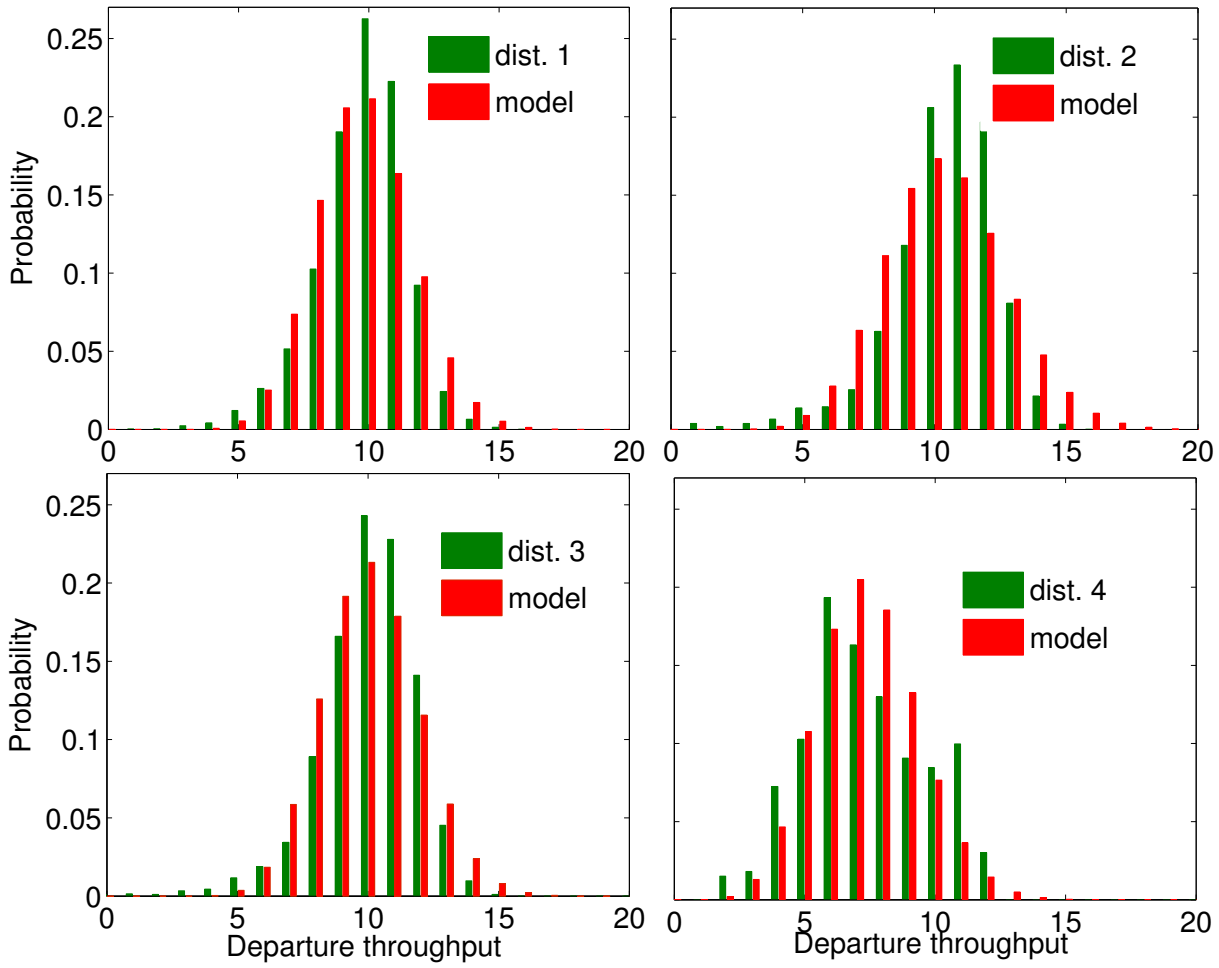


Figure 4-8: Empirical (f_{rw}) and modeled (f_{rm}) probability distributions of the departure throughput of runway configuration 22L |22R under different conditions.

Let μ_1 and μ_2 denote the first and second moment of the empirical distribution f_{rw}. We assume that the service times are generated from an Erlang distribution with parameters $(k, k\mu)$. We estimate these parameters using an approximation based on the method of moments. The output is the Poisson distribution satisfied by the kth arrival of the exponential distribution with rate $(k\mu)$ in a Δ time period that matches the first moment and has the smallest absolute error of the second moment of f_{rw}.

Denote as t_i the time at which the ith service time occurs assuming an infinite queue in the

system. Then we have:

$$\mu_1 = 0 \cdot \mathbb{P}\{t_1 > \Delta\} + \sum_{i=1}^{\infty} i \cdot \mathbb{P}\{(t_i \leq \Delta) \cap (t_i + 1 > \Delta)\} \tag{4.35}$$

$$\mu_2 = 0 \cdot \mathbb{P}\{t_1 > \Delta\} + \sum_{i=1}^{\infty} i^2 \cdot \mathbb{P}\{(t_i \leq \Delta) \cap (t_i + 1 > \Delta)\} \tag{4.36}$$

Given that the times t_i's are generated from an Erlang distribution $(k, k\mu)$, the event of having exactly i services in the time interval Δ has the weighted sum of the probabilities of $(i-1) \cdot k + 1, \ldots, i \cdot k, \ldots (i+1) \cdot k - 1$ of occurrences of a Poisson random variable z with parameter $(k\mu \cdot \Delta)$. Thus:

$$\mu_1 = \sum_{i=0}^{\infty} \left(i \cdot \sum_{j=(i-1)k+1}^{(i+1)k-1} \frac{k - |ik - j|}{k} \cdot e^{(-k\mu \cdot \Delta)} \cdot \frac{(k\mu \cdot \Delta)^j}{j!} \right) \tag{4.37}$$

$$\mu_2 = \sum_{i=0}^{\infty} \left(i^2 \cdot \sum_{j=(i-1)k+1}^{(i+1)k-1} \frac{k - |ik - j|}{k} \cdot e^{(-k\mu \cdot \Delta)} \cdot \frac{(k\mu \cdot \Delta)^j}{j!} \right) \tag{4.38}$$

The method of moments cannot be applied exactly because k is constrained to be a natural number. For this reason we make the following approximation. μ is obtained by numerically solving Equation (4.37) as a function of different increasing k's. For each set of $(\mu, k\mu)$ the error of Equation (4.38) is calculated. When the absolute error increases, we stop the iteration: A further increase in k would imply a further decrease in variance and an increased absolute error in the value of the second moment. For the empirical distribution of the third leaf of the regression tree shown in Figure 4-7b, we obtain the parameters of Erlang distribution (2, 1.38). The mean service time is 2/1.38 min = 1.54 min. The variance of the service time is $2/1.38^2$ min^2 = 1.05 min^2. The corresponding distribution, f_{rm}, of the number of takeoffs in Δ min is depicted in the lower left plot of Figure 4-8. The distributions of the other three leafs are shown in the remaining plots. The four empirical and modeled distributions are similar, as can also be seen in Table 4.1 that compares their standard deviations.

4.4.4 Service time distributions

In this section, we examine the validity of the Erlang-distribution for the service times. The Erlang distribution has been widely used in the air transportation literature (for example [54, 71, 81]) since it was first proposed by Hengsbach and Odoni [56]. As seen in Section 4.4.2, the

Table 4.1: Standard deviation of the distributions f_{rw} and f_{rm}.

Dist.	f_{rw}	f_{rm}
1	2.19	2.30
2	1.86	1.87
3	1.74	1.84
4	2.33	1.95

Erlang distribution offers certain computational advantages, because it can be viewed as a sum of exponential distributions. To the best of our knowledge, validation of the Erlang distribution assumptions has been performed informally with aggregate data, like those of Figure 4-8, or the ones found in earlier work [113]. ASPM data presents several challenges for validating actual service time distributions:

- Aircraft takeoff times are approximated from the ACARS messages and are rounded at the minute [88]. Thus the distribution is discretized, and aircraft can appear to be taking off with zero inter-departure time (from the same runway), if their takeoff time is rounded at the same minute.

- To infer service times from departure times, persistent demand must be guaranteed which requires filtering ASPM data making several assumptions, as was seen in Section 2.3.1. These assumptions, albeit necessary for measuring departure throughput over a 15-minute, may be not be appropriate when applied to estimating service times.

ASDE-X data provides solutions to both these issues. Departure times can be captured with seconds precision. In addition, inter-departure times can be measured conditioned on the actual state of the departure queue, that is the number of aircraft that are *physically* in the queuing area surrounding the threshold of the departure runways. This is enabled by using surveillance data to measure the precise number of aircraft in queue [69, 102].

For these reasons, in this section, we use ASDE-X data. Because ASDE-X data is not available for EWR, we use BOS data from the year 2011. Using ASDE-X data from the year 2011 for BOS major runway configuration 22L, 27 | 22L, 22R, we estimate the instances of persistent demand and fit an Erlang distribution, as described in Section 4.4.3. The empirical and fitted throughput distribution are shown in Figure 4-9. The actual, f_{rw}, and fitted, f_{rm}, distribution parameters are listed in Table 4.2. The parameters $(k, k\mu)$ of the Erlang distribution, g_{rm}, yielding the throughput

Table 4.2: Distributions f_{rw} and f_{rm}.

Distribution	f_{rw}	f_{rm}
Mean	9.81	9.81
Variance	1.91	1.80

distribution f_{rm} are (6, 3.92). The Erlang distribution g_{rm} has an average service time of 1.53 min with variance 0.39 min^2.

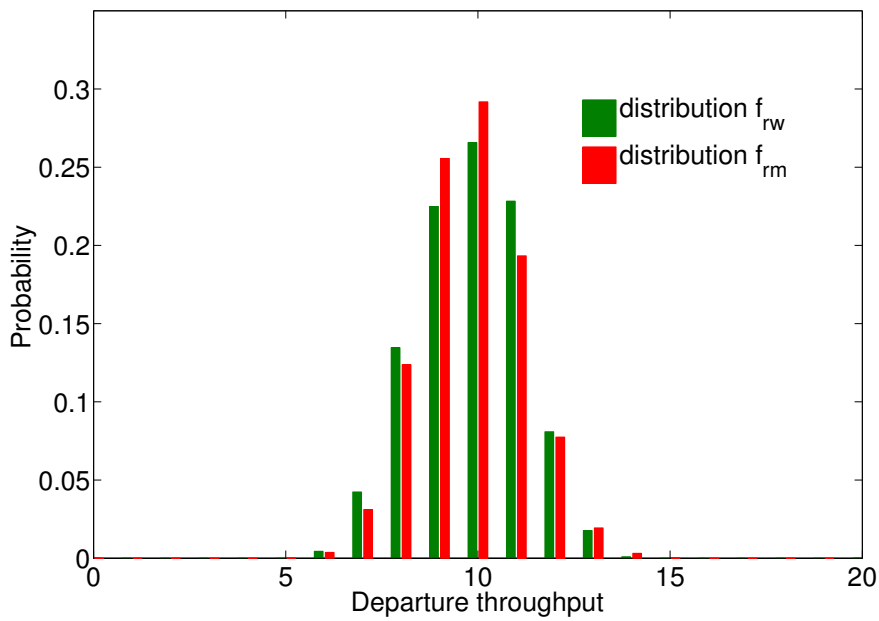


Figure 4-9: Empirical (f_{rw}) and modeled (f_{rm}) probability distributions of the departure throughput of BOS runway configuration 22L, 27 | 22L, 22R.

We turn now to the estimation of service times using ASDE-X data. Conforming to Air Traffic Controllers' phraseology, we define a queue *with pressure* as a departure queue with a sufficient number of aircraft in it that aircraft take off as soon as the runway is available for takeoffs. Conditioning on a queue with pressure, the service time equals the inter-departure time. For estimating the condition that implies a queue *with pressure*, we use the following algorithm, where $d_q(l)$ is defined as the departure queue at the time of takeoff of the lth aircraft . Z is the largest occurrence of $d_q(l)$:

for $i = 1 \rightarrow Z$ **do**

$S_i \leftarrow$ set of inter-departure times of aircraft for which $d_q = i$

end for

```
i ← 1
C ← 0
while C ≠ 1 do
    if inter-departure time distributions of sets S_i, ..., S_n statistically significantly different then
        i ← i + 1
    else
        C ← 1
    end if
end while
return i
```

In other words, the algorithm identifies the value of d_q for which the inter-departure times do not change significantly with the number of aircraft in the queue. For comparing the distributions of service time in sets $S_i, \ldots, S_Z$, we use a non-parametric method, namely the Kruskal-Wallis one-way analysis of variance. For this example, the algorithm returns a value of 5. This result implies that the inter-departure times for this configuration at BOS are distributed differently when there are 4 aircraft in queue and when there are 5 aircraft in queue. We note here, that such high number of aircraft in the queue area is necessary to guarantee that the trailing aircraft is at the runway threshold, and not traveling through the queuing area, or on hold.

For the aircraft in sets $S_5, S_6 \ldots S_Z$, the inter-departure time equals the service time. We note here that the obtained set of service times does not consist of independent samples. Two subsequent takeoffs may be correlated. As shown in Section 2.6.4, a departure of one Heavy aircraft within a 15-minute interval does not impact the departure throughput significantly, because the controllers use the longer separation to perform runway crossings. Thus, the inter-departure time of the surrounding non-Heavy aircraft is shorter than otherwise. In such a scenario, the service times are correlated. Even if there is no Heavy aircraft in the departure queue, a controller might choose to perform a stream of tight non-Heavy departures followed by a stream of runway crossings. In this case, the service-time of the departures will be correlated. This issue has also been considered in the literature, and it is recommended that capacity is defined over a long time period (for instance the *saturation capacity*, the *practical hourly capacity* and the *sustained capacity* are all defined over an hourly-time window [30]). In Chapter 2, the 15-minute period was chosen as good compromise of achieving both a long time period and enough data-points in saturation for the whole length of the time period.

In order to get independent samples of service time distributions, we sample a random set of service times that are all spaced 15 minutes apart. The 15-minute requirement was chosen for consistency with the throughput estimates which are also performed for 15-minute time windows.

The empirical distribution for the service times of aircraft with 5 or more aircraft in the departure queue at the time of their takeoff and which are spaced at least 15 minutes apart is shown in Figure 4-10. From Figure 4-10, we notice that the service time distributions have a very long tail despite having a queue with pressure. We also notice that the distribution's support starts around 50 sec and not 60 sec as theoretically expected. The reason for this is that inter-departure times are measured at the time of wheels-off and not at the start of the takeoff roll where separation is applied by the controllers. The mode of the distribution is at 68 sec. The distribution exhibits also a second distinct pick at around 100 sec which can be attributed to Heavy aircraft departures.

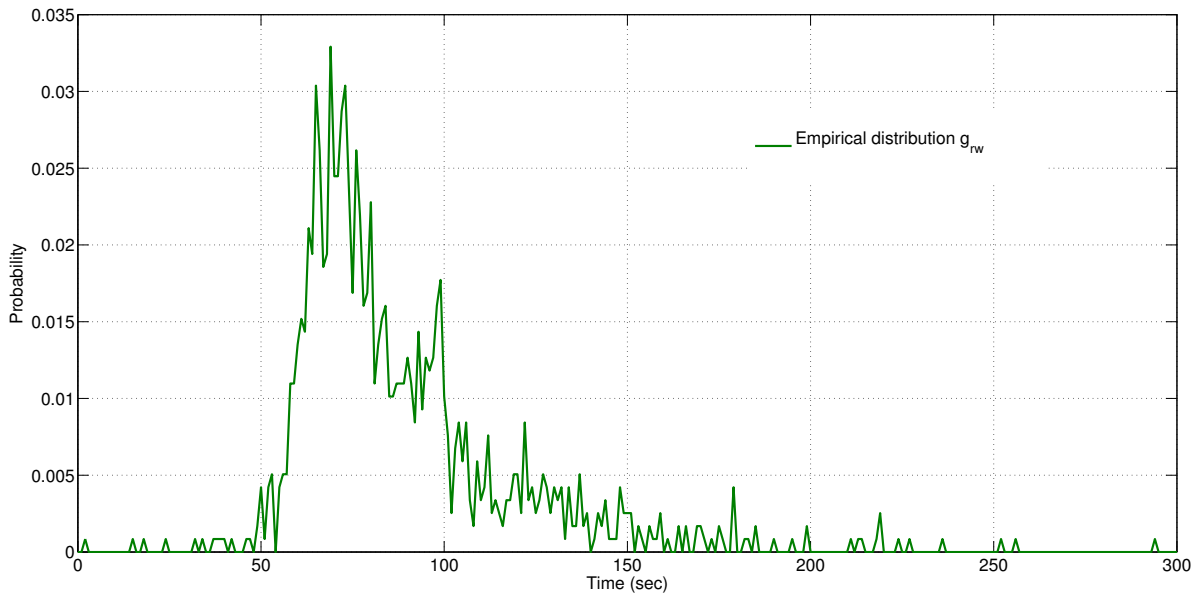


Figure 4-10: Empirical service time probability distribution for departures of runway configuration 22L, 27 | 22L, 22R at BOS.

We compare four different fits to the empirical service time distributions:

1. The maximum likelihood estimation (MLE) Gamma distribution fit (g_{gl}).

   We simply estimate the maximum likelihood parameters of a Gamma distribution fit to the empirical distribution of Figure 4-10.

2. The displaced exponential distribution fit (g_{de}).

   The displaced exponential distribution is given in Equation (4.39). The displaced exponential

distribution is often used in traffic engineering applications because it assumes that there is a minimum headway, d, between vehicles in addition to a probabilistic quantity [123]. Similarly, it can be hypothesized to be a good model for the service time distribution, to model the minimum separation requirement between successive departures.

$$g_{de}(x;\phi,d) = \begin{cases} \phi \cdot e^{-\phi(x-d)} & \text{if } x \geq d \\ 0 & \text{otherwise} \end{cases} \tag{4.39}$$

For fitting the displaced exponential distribution, we choose the parameters (ϕ, d) using the Method of Moments:

$$d + \frac{1}{\phi} = \mathbb{E}[S] \tag{4.40}$$

$$\frac{1}{\phi^2} = var(S) \tag{4.41}$$

3. The Erlang distribution fit from applying an approximate method of moments (MoM), g_{em}. Here, we first use the Method of Moments to fit a gamma distribution to the observed service times histogram. The Method of Moments for the gamma distribution yields estimates $\hat{k}$ and $\hat{\lambda}$ for the shape and scale parameters as follows:

$$\hat{k} = \frac{(\mathbb{E}[S])^2}{var(S)} \tag{4.42}$$

$$\hat{\lambda} = \frac{\mathbb{E}[S]}{var(S)} \tag{4.43}$$

As a next step, we constrain k to be an integer, in order to transform the Gamma distribution to Erlang. We seek to find the Erlang distribution which has the same mean as the observed service time distribution and shape (k) that will result in a variance as close as possible to the observed one:

$$\hat{k} = \lfloor \frac{(\mathbb{E}[S])^2}{var(S)} + 0.5 \rfloor \tag{4.44}$$

$$\hat{\lambda} = \frac{\lfloor \frac{(\mathbb{E}[S])^2}{var(S)} + 0.5 \rfloor}{\mathbb{E}[S]} \tag{4.45}$$

The resulting Erlang distribution will have a mean $\mathbb{E}[L_k]$ and variance $\sigma^2_{L_k}$:

$$\mathbb{E}[L_k] = \frac{\hat{k}}{\hat{\lambda}} = \mathbb{E}[S] \tag{4.46}$$

$$\sigma^2_{L_k} = \frac{\hat{k}}{\hat{\lambda}^2} = \frac{\mathbb{E}[S]^2}{\lfloor \frac{(\mathbb{E}[S])^2}{var(S)} + 0.5 \rfloor} \approx var(S) \tag{4.47}$$

4. The Erlang distribution fit g_{rw}.

   For this fit, we simply use the Erlang distribution g_{rw} with parameters (6, 3.92), which was obtained by fitting f_{rm} to f_{rw}, as seen in Figure 4-9. We note here that f_{rw} comprises all departure throughput observations in saturation, whereas the service time distribution shown in Figure 4-10 comprises independent samples of inter-departure times given a queue with pressure. The objective here is to compare g_{rw} to g_{em}. They both essentially model the same quantity but are estimated differently. We note here that some differences are expected between the two distributions, because the empirical distribution g_{rw} is sampled randomly. A different sampling could yield different parameters.

The results of applying the four different fitting procedures can be seen in Figure 4-11. The estimated parameters for the four distributions can be seen in Table 4.3. From the plots, we note that the displaced exponential fit matches the empirical distribution best. d is estimated to be 0.88 min (53 sec). This means that it captures the minimum separation requirement very accurately. However, it does not predict the mode of the distribution exactly. In addition, it matches the tail of the empirical distribution very well. The Gamma and Erlang fits fail to predict the mode of the distribution, and they overestimate the density of the distribution for values lower than 60 sec. By contrast, they predict the tails of the empirical distribution equally well as the displaced exponential fit. The Gamma and Erlang distributions are different as can be verified from their parameters in Table 4.3. The discrepancy does not result from the approximate method of method applied when deriving distribution g_{em} by rounding $\hat{k}$ (it is rounded to 6 from 5.98) , but from the different method applied (MLE versus MoM). Finally, we notice that distributions g_{em} and g_{rw} are very similar despite the fact that they were derived very differently. This further shows the consistency of the estimated departure throughput *in saturation* and the inter-departure times given a queue *with pressure*.

We conclude that estimating the service time distribution from the throughput distribution, as outlined in Section 4.4.3, calculates accurately not only the mean and the variance of the departure

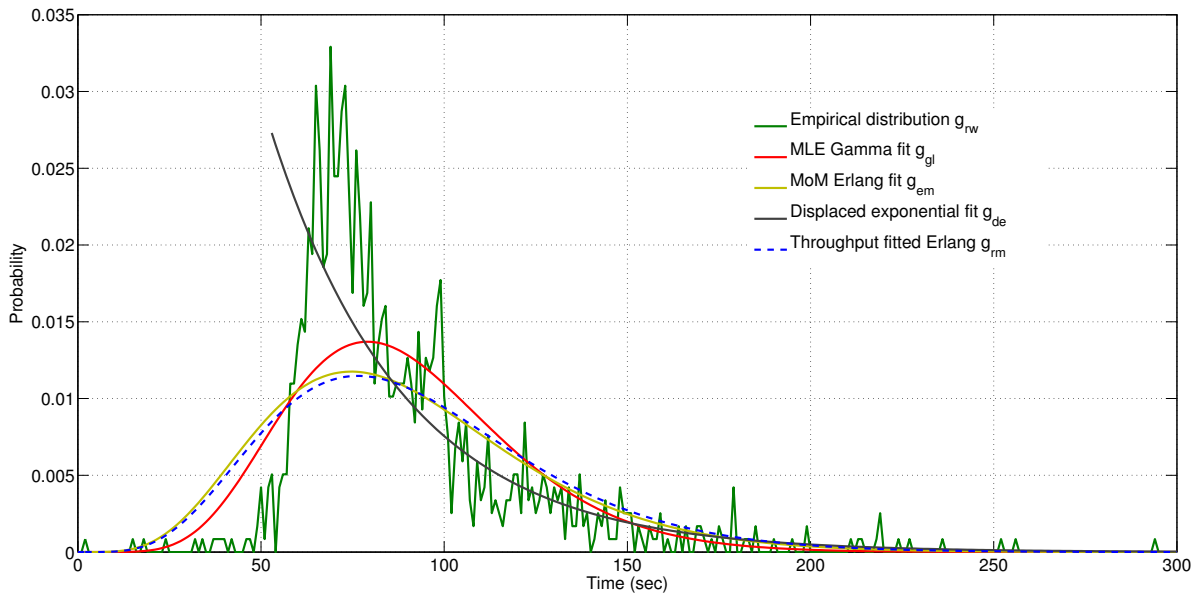


Figure 4-11: Service time probability distribution fits for departures of runway configuration 22L, 27 | 22L, 22R at BOS.

Table 4.3: Distribution parameters

Distribution	Parameter 1 (Shape/ Displacement)	Parameter 2 (Rate)	Mean	Variance
Empirical	–	–	1.49	0.37
g_{gl}	8.54	5.72	1.49	0.26
g_{de}	0.88	0.62	1.49	0.30
g_{em}	6	4.02	1.49	0.37
g_{rw}	6	3.92	1.53	0.39

throughput, but also the mean and variance of the inter-departure time given a queue with pressure. On the other hand, the shape of the fitted Erlang distribution does not fit the actual empirical distribution. From Figure 4-11, it can be hypothesized that the displaced exponential is a better fit. To test this hypothesis, we propose the following method: Use the estimated parameters $(k, k\mu)$ of the fitted f_{rm} distribution to derive a displaced exponential distribution with the same mean and variance:

$$\tilde{g}_{de}(x; \tilde{\phi}, \tilde{d}) = \begin{cases} \tilde{\phi} \cdot e^{-\tilde{\phi}(x-\tilde{d})} & \text{if } x \geq \tilde{d} \\ 0 & \text{otherwise} \end{cases} \tag{4.48}$$

such that:

$$\tilde{d} + \frac{1}{\tilde{\phi}} = \frac{1}{\mu} \tag{4.49}$$

$$\frac{1}{\tilde{\phi}^2} = \frac{1}{k\mu^2} \tag{4.50}$$

For this example the parameters are calculated (0.91, 0.62). As expected, the parameters are very similar to those of g_{de} (Table 4.3). The corresponding $\tilde{f}_{de}$ is shown in Figure 4-12 and matches well with the empirical distribution. We also note from Table 4.4 that it has smaller KL divergence from the actual throughput distribution compared to f_{rm}. Thus, we conjecture that distribution $\tilde{g}_{de}$ accurately represents the service time distribution. It models the minimum service time requirement, the observed tail of the empirical service time distribution and the associated departure throughput distribution.

We model the service times in our model with Erlang distributions as explained in Section 4.4.2. However, we will perform sensitivity analysis of the results. For this, we will use service time distributions given by Equation (4.48) to investigate if they lead to different queuing delays than those calculated using the Erlang distribution with parameters $(k, k\mu)$.

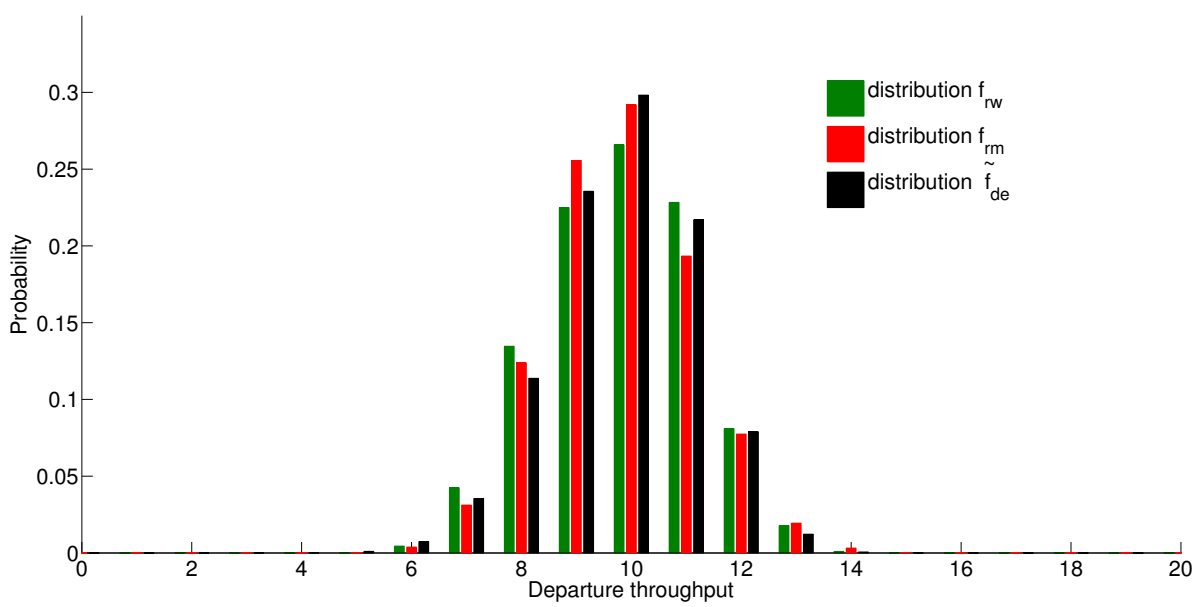


Figure 4-12: Empirical throughput distribution, f_{rw}, and fits f_{rm} and $\tilde{f}_{de}$ for departures of BOS runway configuration 22L, 27 | 22L, 22R.

Table 4.4: Comparison of the distributions f_{rw}, f_{rm} and f_{de}.

Distribution	Mean	Variance	KL Divergence
f_{rw}	9.81	1.90	–
f_{rm}	9.81	1.80	0.0102
$\tilde{f}_{de}$	9.81	1.77	0.0079

The need for sampling

Here, we demonstrate the need to sample the service times with an example. In Figure 4-10, we show the empirical distribution g_{rw} of the sampled service times given a queue with pressure. Suppose we have only this distribution and we generate the corresponding throughput distribution, f_{sf} in a 15-minute period. We consider now the distribution g_{sa} of all the service times given a queue with pressure. We similarly generate the corresponding throughput distribution, f_{sa}, in a 15-minute period. We compare the distributions f_{sf} and f_{sa} to the empirical distribution f_{rw} in Figure 4-13. We also compare the three distributions in Table 4.5. We notice that f_{sa} is a less variable throughput distribution than f_{rw}. This results from constructing the empirical distribution f_{sa} using dependent observations. From Table 4.5 we also note that f_{sa} has a higher KL-distance from f_{rw} than f_{sf} has.

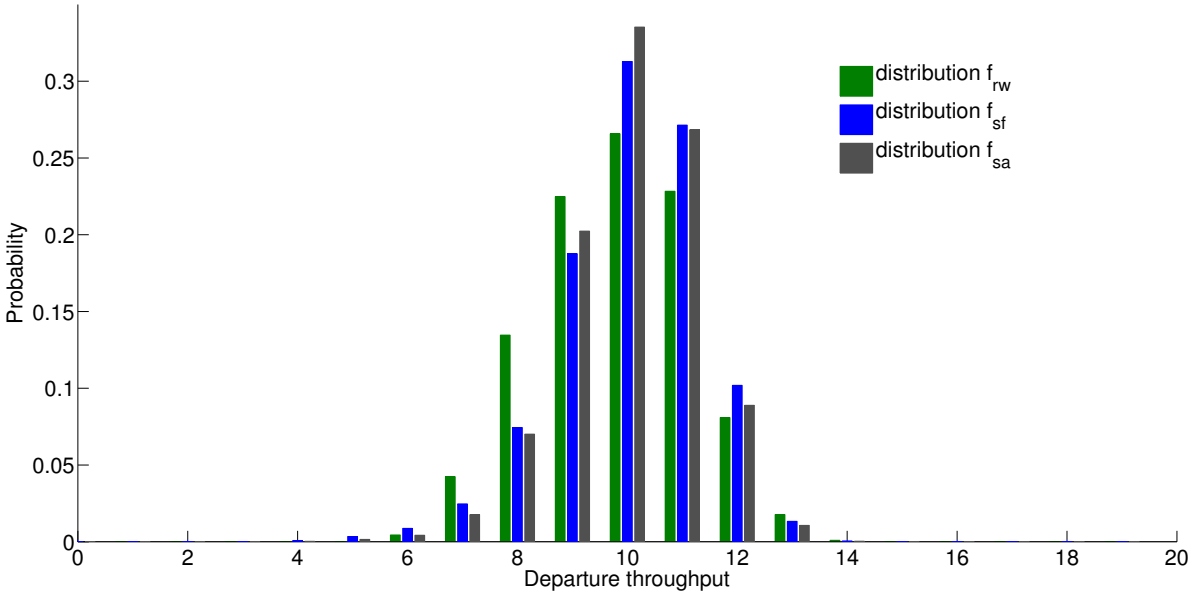


Figure 4-13: Empirical throughput distribution, f_{rw}, and fits f_{sf} and f_{sa} for departures of BOS runway configuration 22L, 27 | 22L, 22R.

Table 4.5: Comparison of the distributions f_{rw}, f_{sf} and f_{sa}.

Distribution	Mean	Variance	KL Divergence
f_{rw}	9.81	1.90	–
f_{sf}	10.05	1.78	0.0445
f_{sa}	10.06	1.49	0.1053

Service time and fleet mix

The empirical service time distributions can be further parametrized by fleet mix. The Heavy jets are theoretically expected to have a service time of at least 2 minutes (120 sec). In Figure 4-14, we show the empirical service time distributions parametrized by the type of aircraft taking off. The distribution of the service times for the Heavy aircraft is clearly distinct from that of the non-Heavy jets. Heavy jets have much longer service times, as it was expected given their separation requirement. The mode of the distribution is at 105 sec. The service time is measured as the difference between successive wheels-off times, and thus it is on average shorter than the separation at the start of the takeoff roll if the leading aircraft is a Heavy. Heavy aircraft have on average longer roll times. The average service time of non-Heavy aircraft is 87 sec and of Heavy aircraft 119 sec.

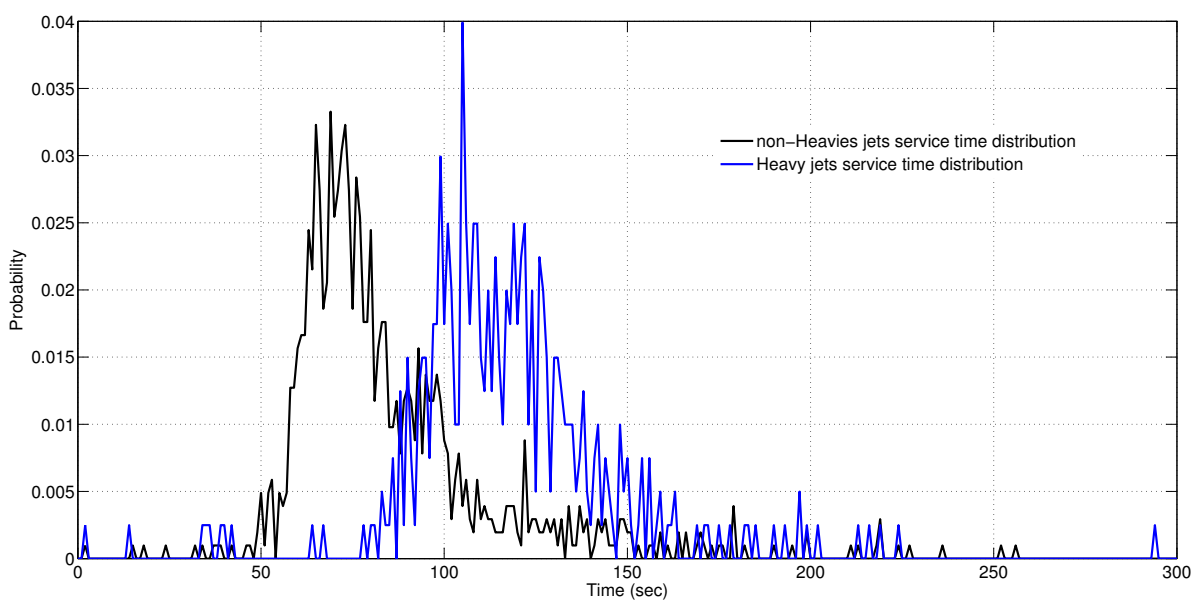


Figure 4-14: Empirical service time probability distributions for departures of BOS runway configuration 22L, 27 | 22L, 22R for Heavy and non-Heavy jets.

Finally, we discuss how these findings compare to the findings regarding the jet departure capacity of Section 2.6.4. There, we concluded that the departure throughput does not depend on Heavy aircraft, whereas from Figure 4-14, we clearly see that Heavy aircraft are separated from the subsequent departures on average for much longer than non-Heavy aircraft (32 sec). However, Figure 4-14 does not show how departing Heavy aircraft impact the service times of surrounding non-Heavy aircraft. This is precisely the value of the parametrized capacity envelope. Figure 2-14b conveys that controllers utilize the runway in such way that the impact of Heavy aircraft is diminished in a 15-minute time window. Regarding Figure 4-14, this would imply that short non-Heavy aircraft service times are correlated with Heavy departures in the surrounding time-window. Similarly, the impact of an arrival bank will not be revealed on the inter-departure time of a single aircraft, but will be shown in the 15-minute departure throughput.

To fully capture the separation requirements and the impact of the exogenous variables (arrivals crossings, route availability, props fanning), one would need a hidden Markov model where the exogenous variables (arrivals crossings, route availability, props fanning), along with the endogenous one (Heavy vs. non-Heavy) would explain the separation times. The complexity of such model renders it impractical for both modeling and simulation. For this reason, we estimate the departure throughput distribution given the endogenous and exogenous conditions in each 15-minute time-window, and allocate service-time distributions to individual aircraft consistent with the estimated departure throughput distribution.

4.4.5 Module 2 outputs

Module 2 takes a runway schedule (RS) as input from Module 1. For each 15-minute time period, i, given the parameters of the service time distribution $(k_i, k\mu_i)$, and the inter-arrival times c_l, it calculates the queue state probability vector $\mathbf{p}_{Q_l}$. For each aircraft l it calculates its effective queue length $\tilde{q}_l(j)$ and its effective queuing delay $\tilde{d}_l(j)$. The takeoff time of each aircraft is assumed to be $C_l + \tilde{d}_l(j)$, that is, the sum of its arrival time at the departure queue and is effective queuing delay. Repeating this process for each aircraft yields the expected takeoff schedule. From the runway schedule, and the expected takeoff schedule we calculate airport performance characteristics such as the expected throughput, expected queue length, and the expected number of aircraft on the ground.

We also note that Module 2 is independent from Module 1. Thus, if we have the runway schedule of two departure runways, we can solve Equation (4.18) independently for each of them to calculate

their performance characteristics. As already explained, this framework was applied in other work for modeling the departure process from runways of 21R and 22L of DTW [76].

4.5 Model results for EWR

In this section, we discuss the prediction results for the most frequently used runway configuration of EWR, (VMC; 22L | 22R), in 2011. The unimpeded taxi-out times are estimated using ASPM and Flightstats data from the year 2011, as explained in Section 4.3.1. Similarly, the throughput distributions are estimated from ASPM, Flightstats and RAPT data, as explained in Section 4.4.3, and are shown in Figure 4-8. As noted in Section 4.3.2, α is calculated so that the predicted median taxi-out time equals the actual median time (18 min) and equals 0.27 min/AC.

Figure 4-15 shows the frequency of the different congestion states observed in the operational data and predicted by the model. The model predicts the airport being at all congestion levels as often as observed. Figure 4-15 also shows the expected taxi-out time as a function of the number of aircraft taxiing-out at the time of pushback for both the actual operations and the modeled operations. We note that the taxi-out time is predicted very accurately for all traffic conditions. Table 4.6 contains more detailed statistics about the number of aircraft and the taxi times in different congestion levels. In agreement with the plots of Figure 4-15, the model predicts accurately both the frequency of the different congestion states and the taxi-out times at each state. We note, in particular, that the predicted number of flights at congestion states greater than or equal to 15 is only 2% higher than that in actual operations, and their predicted mean taxi-out time is over-estimated only by 2%.

Table 4.6: Aggregate taxi time predictions for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Model mean taxi time (min)
all	65990	20.41	65977	20.19
$(N \leq 8)$	27387	15.89	27964	15.10
$(9 < N \leq 14)$	22594	19.51	21683	19.47
$(N \geq 15)$	16009	29.42	16330	29.84

Figure 4-16 shows the predicted throughput of configuration 22L | 22R at EWR in 2011 as a function of the congestion state N. The model predicts both the mean throughput and the median throughput very accurately in all traffic conditions. However, the model is not capable of producing

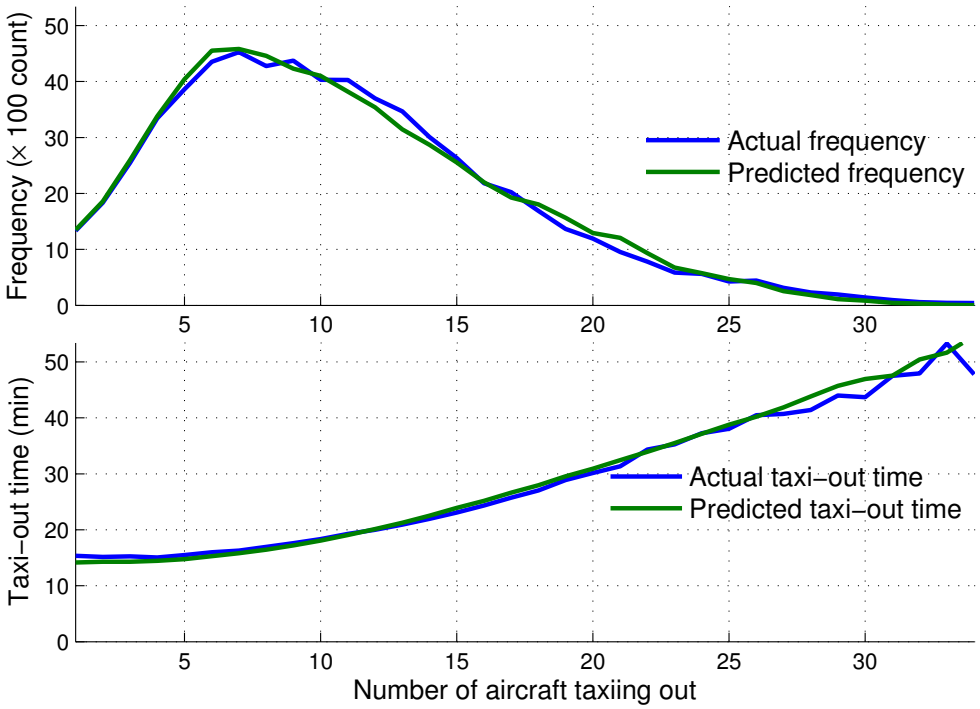


Figure 4-15: Actual and modeled frequency of each state N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 22L | 22R in year 2011.

estimates of the variance of the departure throughput.

We also list the prediction results for both the congestion state (N) and the predicted 15-minute throughput at each minute at which there was traffic on the ground in the actual data ($N(t) > 0$) in Table 4.14. We calculate the mean error (ME), the mean absolute error (MAE) and the root mean square error (RMSE). As far as the predictions of the congestion state are concerned, we note that the negative mean error of -0.2 indicates that the state is underpredicted by 0.2 units (AC taxiing out) on average. The condition $N(t) > 0$ is applied for removing the data-points of no traffic, which are of no interest. We also list the results for the conditions when $N(t) > 10$, because these are the higher congestion states, and they are more relevant to the problem of congestion. We note that given $N(t) > 10$, the congestion state is under-predicted by 0.9 units. This is due to the fact that very high congestion states are also related to rare events (safety incidents, temporary runway closures, mechanical problems, constraints at the destination airports, etc.), which are not modeled.

The predictions for the 15-minute throughput at each minute are also listed in Table 4.7. The errors are lower than the errors of the congestion state predictions because the average throughput

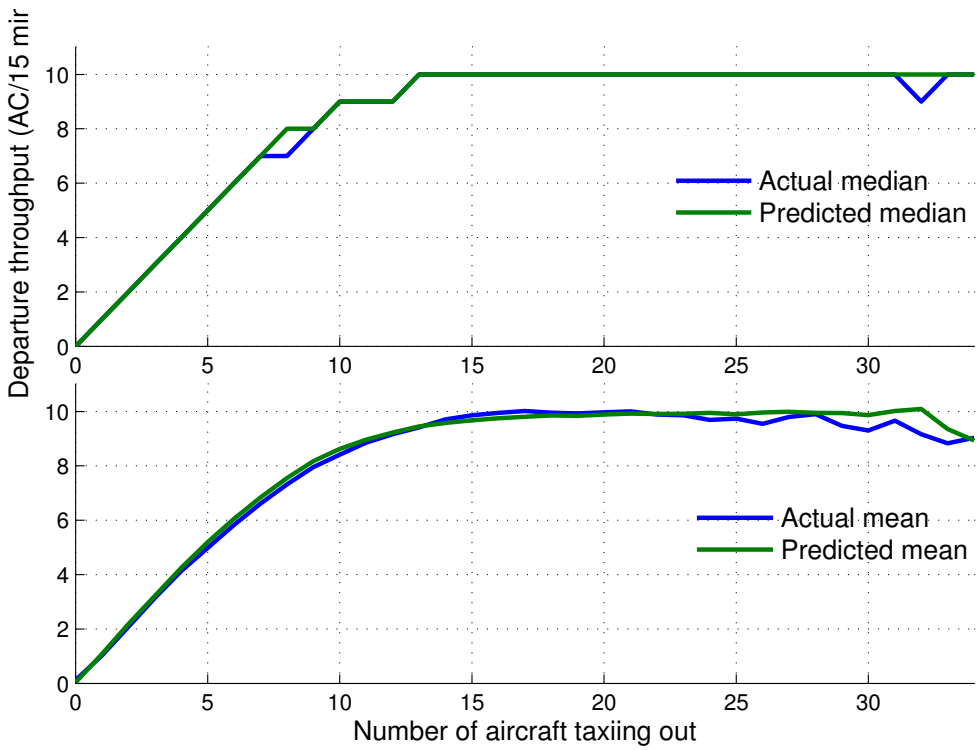


Figure 4-16: Actual and modeled throughput of all states N for EWR runway configuration 22L | 22R in year 2011: Mean (top); Median (bottom).

is not very sensitive to the exact value of the congestion state $N(t)$. Neighboring values of N may imply the similar throughput, or the same throughput if they are in saturation. We note here that the throughput comparison is performed by using time as a basis. At each minute t, the model predicts both the number of aircraft on the ground and the throughput at this minute t with a mean absolute error of 1.7 AC and 1.1 AC/15 min respectively.

Table 4.7: Prediction statistics for the congestion state and the throughput for EWR runway configuration 22L | 22R in year 2011.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.20	1.71	3.03	-0.91	3.00	4.5
Throughput (AC/15 min)	0.00	1.14	1.60	-0.32	1.37	1.85

4.5.1 Comparison to a deterministic model

In this section, we compare the results of the model to an alternative model where the service times are deterministic. In this setting, each aircraft is served deterministically with the average service

time of the time-window i in which it takes off . We use the same runway schedule as that used in the previous section and compare the results of the two models.

In Table 4.8, we list the results for the taxi-out time predictions using the deterministic model. It is evident that the deterministic model underestimates the average taxi-out time. It overestimates the number of aircraft that take-off under light congestion ($N \leq 8$) and analogously it underestimates the number of aircraft that take-off in higher congestion states ($N \geq 15$). However, for the aircraft that are predicted to take-off in higher congestion, their taxi-out time is predicted very accurately. It is only 0.5% lower than the actual one.

In high congestion, the deterministic and stochastic model converge to the same delays [73]. However, the deterministic model transitions to higher-congestion states less often than the stochastic one. Thus, fewer flights end up facing a long queue in the deterministic than in the stochastic model. For the ones that end up facing a long queue in both models, their delays are very similar.

Table 4.8: Aggregate taxi time predictions using a deterministic model for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Model mean taxi time (min)
all	65,990	20.41	65,977	19.10
($N \leq 8$)	27,387	15.89	30,277	14.77
($9 < N \leq 14$)	22,594	19.51	21,494	18.47
($N \geq 15$)	16,009	29.42	14,204	29.27

In Figure 4-17, we show the taxi-out time of each aircraft as a function of its takeoff queue, that is the number of aircraft that take-off between its pushback and takeoff time. At low values of the takeoff queue, ($N_{toff} \leq 6$) both the deterministic and the stochastic model overestimate the taxi-out time for a given value of the takeoff queue compared to the actual data. This is because of the selection-bias in the actual data, as discussed in Section 4.3.1. The flights with shorter unimpeded taxi-out times than the average will be associated with lower take-off queues. In both the deterministic and stochastic model, flights are assumed to travel with their *average* unimpeded taxi-out time to the runway. Thus the measurements are not biased and their curves lie higher than that of the actual data.

For intermediate values of the takeoff queue ($8 < N_{toff} < 18$), the stochastic model predicts the same average taxi-out time as it was observed. By contrast, the deterministic model consistently underestimates the taxi-out time. This discrepancy can help explain the necessity for a stochastic model. The takeoff queue of an aircraft is distributed over its route from gate to runway. At

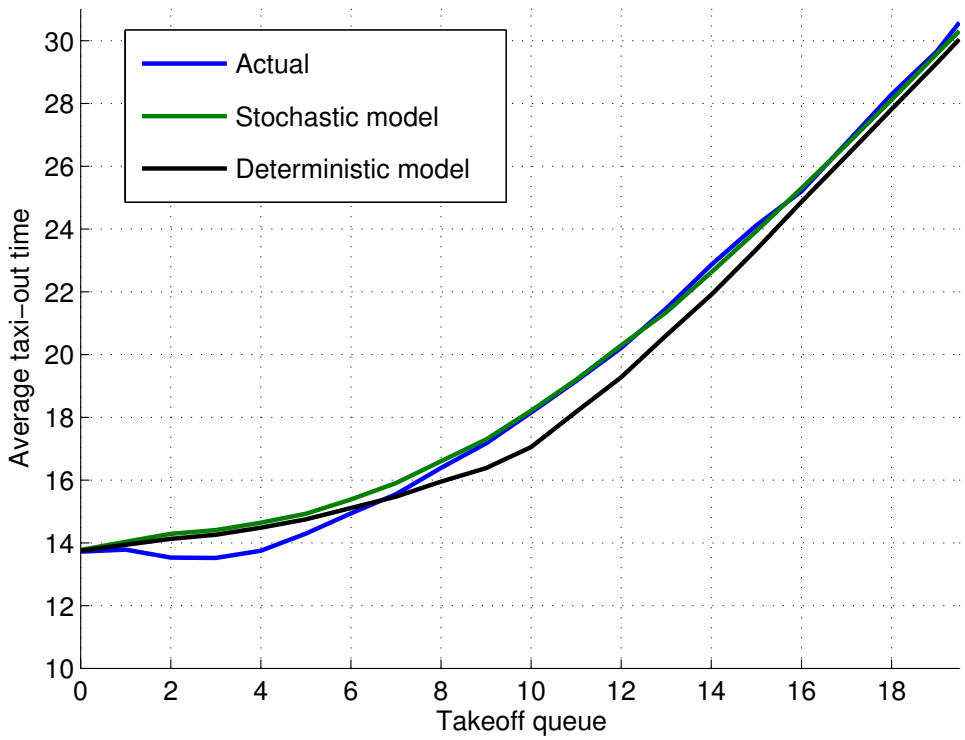


Figure 4-17: Actual and predicted average taxi-out time as a function of the takeoff queue of each aircraft.

medium values of the takeoff queue, the departure queue has a positive probability of being empty. In these cases, an event of service time shorter than the average will not benefit other aircraft: The departure queue is empty, thus no aircraft can take advantage of the runway availability. By contrast, an event of service time longer than the average has a positive probability of inducing a delay cost on other aircraft taxiing out. This holds even given an empty queue at the time of departure. The longer service time has a positive probability of resulting in a queuing time for the next aircraft that reaches the runway and that would have arrived at an empty system in the case of a deterministic service process.

As the values of the takeoff queue increase, the probability of having a positive departure queue at all times increases. In this case, a shorter service time of an aircraft translates to a shorter queuing time for the aircraft in queue. Given a queue *with pressure*, takeoffs take place at the same rate as the service rate and on average each aircraft experiences a delay proportional to the number of aircraft getting served while it taxies-out (its takeoff queue) independently of the service process characteristics. Thus, as the takeoff queue increases, the predicted taxi-out time of both models converge, as shown in Figure 4-17.

4.5.2 Comparison to Monte Carlo simulations

The stochastic model described so far showed promising results. We examine the sensitivity of the model to two important assumptions that were used in its construction.

- The stochasticity of the travel times can be ignored, thus the travel times and the runway schedule can be assumed to be deterministic. We can derive the runway schedule using the expected travel time (Equation 4.9).

- The distribution of the service times can be assumed Erlang although this is a bad fit to the actual service time distributions (Figure 4-11).

To test the robustness of the model to these assumptions, we run 100 Monte-Carlo simulations of four different simulation settings, relaxing each assumption sequentially. The four different simulation settings are described as follows:

1. Fixed unimpeded taxi-out times (equal to their expected value), and service times sampled from the Erlang distribution. This simulation uses the same assumptions as the model, but simulates the Erlang service process instead of modeling it. It provides a benchmark for the statistical significance of the errors of the other Monte Carlo simulations. For an increasing number of trials, it should yield, on average, the same results as the model.

2. Fixed unimpeded taxi-out times (equal to their expected value), and service times at each time-window sampled from the displaced exponential distribution with the same mean and variance as the Erlang distribution (Equation 4.48). This simulation will reveal the error from using an Erlang distribution for the service times.

3. Unimpeded taxi-out times sampled from their distribution, and service times sampled from the Erlang distribution. In each run, the unimpeded taxi-out time of each aircraft is sampled from the Lognormal distribution that describes the unimpeded taxi-out time of its airline. This simulation will reveal the error of assuming fixed values for the unimpeded taxi-out times.

4. Unimpeded taxi-out times sampled from their distribution, and service times at each time-window sampled from the displaced exponential with the same mean and variance as the Erlang distribution (Equation 4.48). This simulation will reveal the combined error from

using an Erlang distribution for the service times and assuming deterministic unimpeded taxi-out times.

In Tables 4.9-4.12 , we list the results for the four Monte Carlo simulation settings. From comparing the results of simulation settings 1 and 2, and simulation settings 3 and 4 pairwise, we notice, that the estimated taxi-out times are not sensitive to the distribution family assumed for the service times. Comparing the results of simulation settings 3 and 4 to those of simulation setting 1 and the model we notice that, in agreement with earlier research [107], the taxi-out time estimates show very little sensitivity to the assumption of fixed travel times. In particular, the average taxi-out time increases only from 20.19 min to 20.25 min. We conclude that the delays estimations is fairly insensitive to the two assumptions of the model.

We also note that the average taxi-out time predicted by the first Monte Carlo simulation setting matches the expected taxi-out time calculated with the model. This provides a numerical validation for the developed analytical model. However, the congestion state (N) frequencies predicted by the Monte Carlo simulation as seen in Table 4.9 do not match the frequencies predicted by the stochastic model and listed in Table 4.6. This is because of an approximation used, by construction, in the model that was explained in Section 4.4.5. The congestion states N are estimated from the expected runway schedule, that is, from the expected takeoff times. The results of the model would match the results of the Monte Carlo simulation, if we calculated the expected congestion state at each minute using the queuing state probability vector.

Finally, we note here that the Monte Carlo simulation is useful for estimating the distribution of taxi-out times. Figure 4-18 shows the histogram of the measured taxi-out times and the simulated taxi-out times. For the simulation, we simply use a single run of simulation setting 4. We notice that the simulation yields very good estimates of the frequency of each observed value for taxi-out times lower than 30 minutes. This provides further evidence of the unimpeded taxi-out time distributions are estimated correctly. For taxi times higher than 30 min, the fit is less good, presumably because of the impact of unmodeled factors.

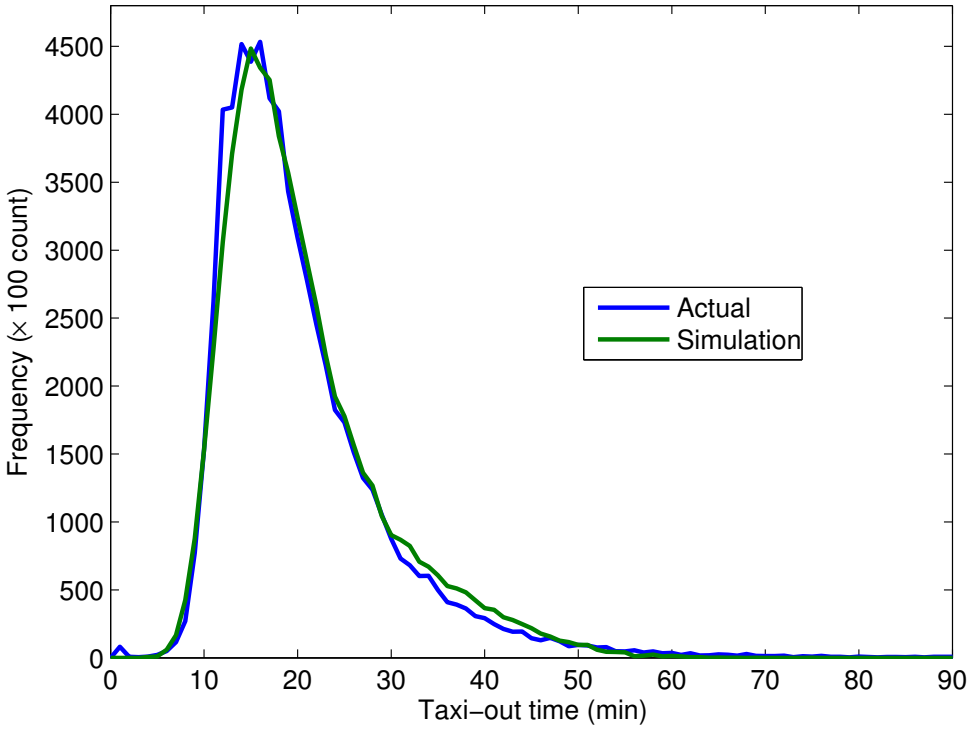


Figure 4-18: Actual and modeled histogram of taxi-out times for EWR runway configuration 22L | 22R in year 2011.

Table 4.9: Averaged taxi time predictions from Monte Carlo simulations of simulation setting 1 for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Simulated mean taxi time (min)
all	65990	20.41	65970	20.18
$(N \leq 8)$	27387	15.89	29095	15.05
$(9 < N \leq 14)$	22594	19.51	20966	19.19
$(N \geq 15)$	16009	29.42	15909	30.87

Table 4.10: Averaged taxi time predictions from Monte Carlo simulations of simulation setting 2 for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Simulated mean taxi time (min)
all	65990	20.41	65969	20.18
$(N \leq 8)$	27387	15.89	29109	15.06
$(9 < N \leq 14)$	22594	19.51	21010	19.17
$(N \geq 15)$	16009	29.42	15850	30.93

Table 4.11: Averaged taxi time predictions from Monte Carlo simulations of simulation setting 3 for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Simulated mean taxi time (min)
all	65990	20.41	65955	20.25
$(N \leq 8)$	27387	15.89	28887	15.21
$(9 < N \leq 14)$	22594	19.51	21055	19.29
$(N \geq 15)$	16009	29.42	16013	30.59

Table 4.12: Averaged taxi time predictions from Monte Carlo simulations of simulation setting 4 for EWR runway configuration 22L | 22R in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Simulated mean taxi time (min)
all	65990	20.41	65955	20.20
$(N \leq 8)$	27387	15.89	28909	15.22
$(9 < N \leq 14)$	22594	19.51	21280	19.24
$(N \geq 15)$	16009	29.42	15766	30.64

4.6 Predictive ability of the proposed model

In this section, we assess the capability of the model to predict operations in different years. For this, we predict taxi-out times in the years 2010 and 2007. We use the realized pushback schedule, the arrival throughput and the $RAPT$ value for the times that runway configuration 22L | 22R was in use as an input to the developed model. We do not have $RAPT$ data for the year 2007, and thus we expect our predictions to be worse for this year, compared to year 2010.

4.6.1 Predictions for EWR in year 2010

Here, we validate the predictions of the model by applying it for predicting the operations for this runway configuration in 2010. In Figure 4-19, we show the actual and predicted frequency of each congestion state and the predicted average taxi-out time as a function of the number of aircraft taxiing out at the time of pushback. We observe that, qualitatively, the predictions match the actual operations as well as they did for 2011. Table 4.13 lists the predicted average taxi-out times at different states. We notice that the number and the delays of the flights in higher congestion states ($N \geq 15$) is underestimated. The median taxi-out time is underestimated as well: It is measured 18 min, but its estimated value is 17.94 min.

Table 4.13: Aggregate taxi time predictions for EWR runway configuration 22L | 22R in year 2010.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Model mean taxi time (min)
all	63633	20.83	63585	20.30
($N \leq 8$)	27945	15.46	27942	15.19
($9 < N \leq 14$)	19530	19.49	20234	19.69
($N \geq 15$)	16158	31.73	15409	30.32

Figure 4-20 shows the predicted throughput of segment (VMC; 22L | 22R) at EWR in 2010 as a function of the number of departing aircraft on the ground. As can be observed, the model overpredicts both the mean throughput and the median throughput in high congestion states. From the actual mean and median throughput curves shown in Figure 4-20, we note that the throughput is 2010 decreases for congestion states higher than 25. This decrease is not predicted by the model, and so the taxi-out times are underestimated.

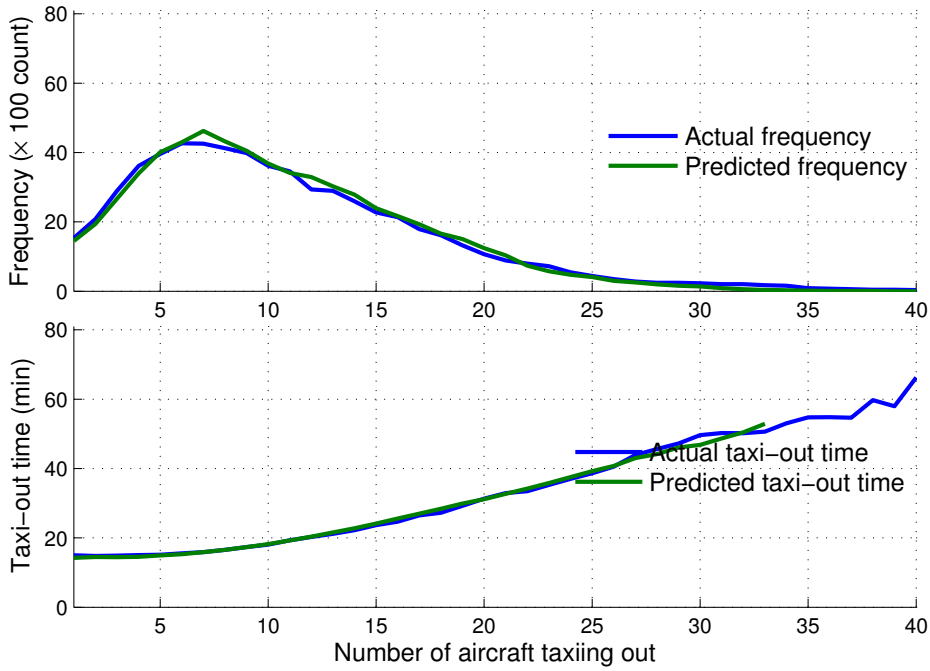


Figure 4-19: Actual and modeled frequency of all states N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 22L | 22R in year 2010.

Table 4.14: Prediction statistics for the congestion state and the throughput for EWR runway configuration 22L | 22R in year 2010.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.33	1.78	3.49	-1.55	3.45	5.56
Throughput (AC/15 min)	-0.01	1.12	1.61	-0.34	1.41	1.92

4.6.2 Predictions of delays on individual days

Given the times at which flights push back, we would like to estimate their travel time to the runway, the amount of time that they spend in the departure queue, the overall state of the airport surface (for example, the number of departures on the ground), and the length of the departure queue. The research question is whether the model, the parameters of which were estimated with operational data from 2011, can predict operations on a day of 2010 that runway configuration 22L | 22R is in use.

Figures 4-21, 4-22 and 4-23 show the results of making predictions using the pushback schedule for a 13-hour period on three days in EWR, together with the observed data. The upper plot shows

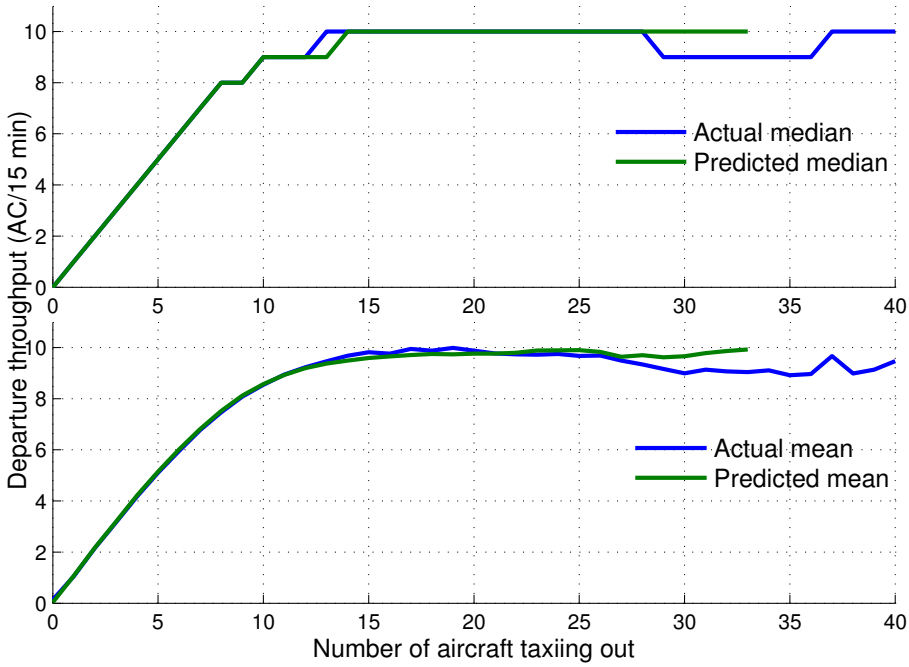


Figure 4-20: Actual and modeled throughput of all states N for EWR runway configuration 22L | 22R in year 2010: Mean (top); Median (bottom).

the observed and predicted number of departures in a 15-minute window, the middle plot contains the average taxi-out times of the flights that push back in the corresponding 15-minute window, and the lower plot show the average predicted departure queue size for each 15-minute window and the number of pushbacks in this 15-minute window for that day. For the average taxi-out time predictions we use the estimated variance of the queuing delays to provide a confidence interval for the estimates. The dashed lines are calculated from the standard deviation of the queuing delay for the flight that is expected to takeoff in the middle of each 15 minute interval. The queuing delays of individual flights are clearly correlated, thus the variance of the queuing delay of a representative flight is a simple measure of the variability of the delays in each time-window.

In Figure 4-21, we notice that the departure queue is expected to comprise at least 10 AC for 3.5 hours, between 1745 hours and 2115 hours. The persistently long queue induces an increased variance of the queuing delays as can be seen from the middle plot of Figure 4-21. For flights that pushed back between 1945 and 2045 hours, the queuing delay standard deviation is around 25 minutes. Concerning the average taxi-out time prediction, we note that it is underestimated between 1600 and 1800 hours and overestimated between 1930 and 2030 hours. It is clearly very hard to accurately predict taxi-out times without updates in such a dynamic and stochastic system

where the queue remains under pressure for a long time.

The throughput is predicted very accurately. During the very busy hours (1800 to 2100 hours), the error is at most 2 AC/15 min. However, each error propagates to the taxi-out times of all later flights until the end of the day, since the queue never empties. The standard deviation of the delays, albeit very approximate, provides potentially helpful information to the system operators on the delay risks of such dense pushback schedule.

Figures 4-22 and 4-23 offer similar insights. Figure 4-22 shows a day with much less demand. We notice that both the average delays and their variability is much smaller than those of Figure 4-21. Figure 4-23 is a day with high demand, but not-uniformly distributed during the day. At 1200 hours, the taxi-out time is both in the predictions and the actual data very long (30 minutes). However, the departure push is very short, the queue does not build up and the delay queuing variance stays very small. By contrast, for the time period between 1730 and 1815 hours shorter expected delays are predicted. However, the predictions are much more uncertain.

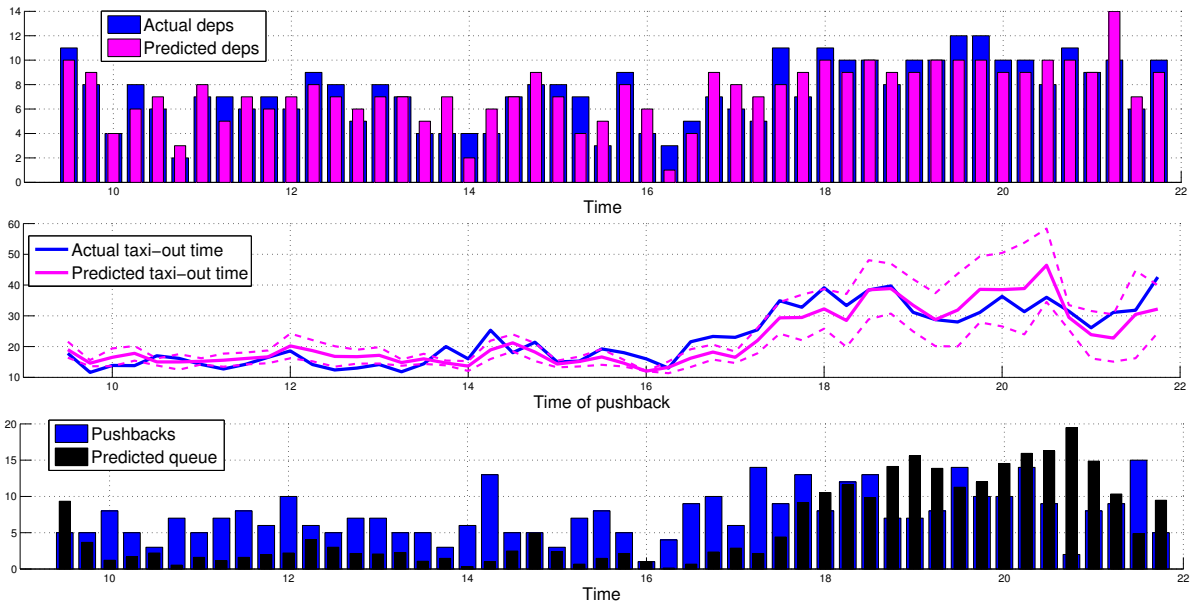


Figure 4-21: Predictions of departure throughput, average taxi-out times and departure queue lengths in each 15-min interval over a 13-hour period on Thursday, August 5, 2010.

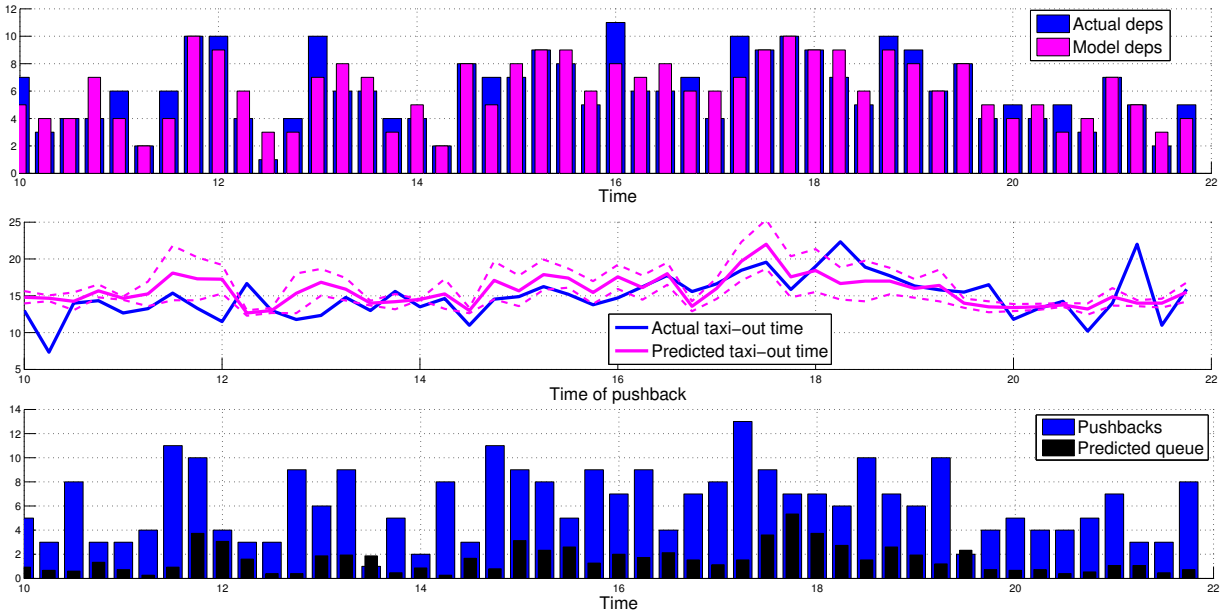


Figure 4-22: Predictions of departure throughput, average taxi-out times and departure queue lengths in each 15-min interval over a 12-hour period on Friday, 26 November, 2010.

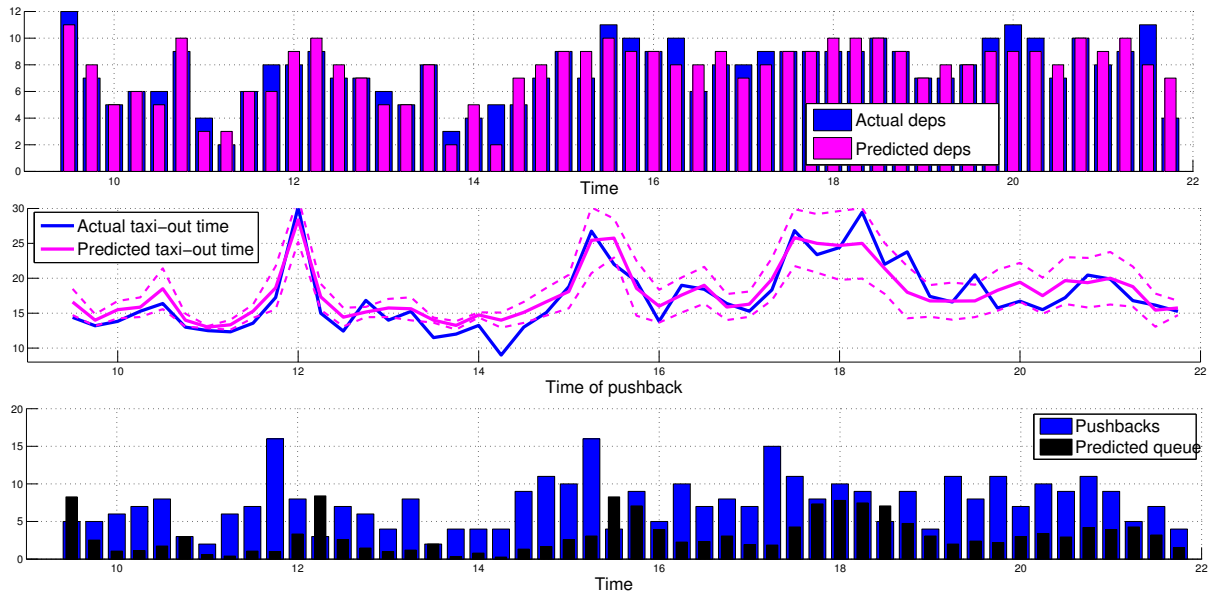


Figure 4-23: Predictions of departure throughput, average taxi-out times and departure queue lengths in each 15-min interval over a 13-hour period on Wednesday, 8 December, 2010.

4.6.3 Predictions for EWR in year 2007

We also use the model (developed with 2011 data) to predict congestion and delays at runway configuration 22L | 22R in 2007, the worst recent year in terms of delays. In Figure 4-24, we show the actual and predicted frequency of each congestion state, and the predicted average taxi-out time as a function of the number of aircraft taxiing out at the time of pushback. We observe that the quality of the predictions deteriorates. The model still predicts a much higher average taxi-out time than in 2010 or 2011. In particular, from Table 4.15, we observe that it underestimates the number of flights that have average taxi-out times of 36 min only by 15% (18640 flights compared to 22101, in real data). Thus, despite the underestimation of the average delays, the model provides useful information for the magnitude and severity of delays expected as a result of the very high demand in year 2007.

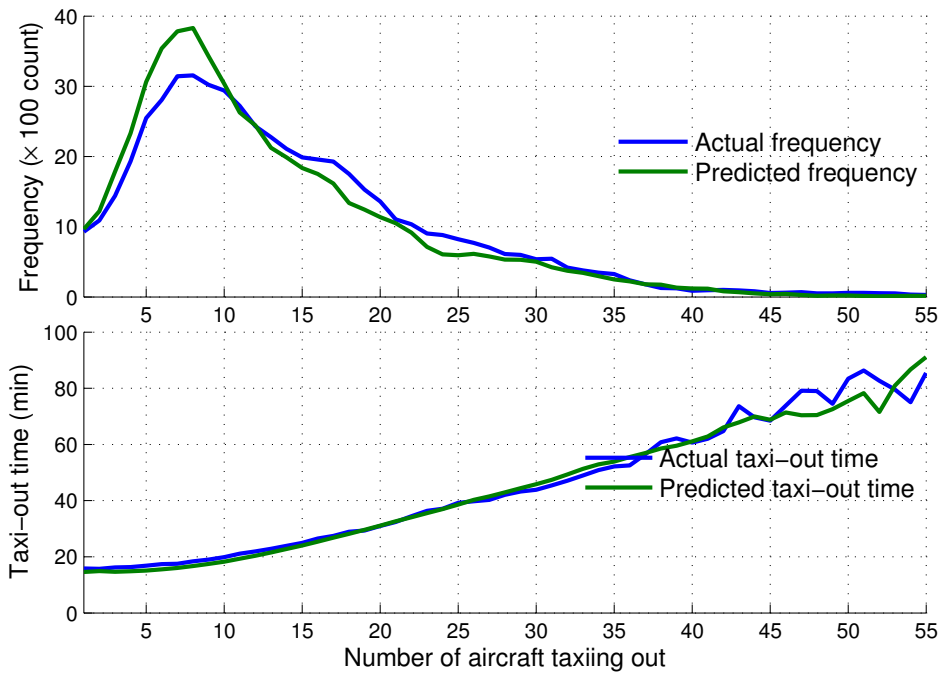


Figure 4-24: Actual and modeled frequency of all states N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 22L | 22R in year 2007.

Figure 4-25 shows the predicted throughput of segment (VMC; 22L | 22R) at EWR in 2007 as a function of the number of departing aircraft on the ground. The model overpredicts both the mean throughput and the median throughput in intermediate congestion states. Thus, flights are predicted to take-off at a higher rate than they actually do and delays are underestimated.

Table 4.15: Aggregate taxi time predictions for EWR runway configuration 22L | 22R in year 2007.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Model # of flights	Model mean taxi time (min)
all	55506	25.83	55704	23.48
$(N \leq 8)$	17899	17.09	21424	15.49
$(9 < N \leq 14)$	15506	19.49	15640	19.61
$(N \geq 15)$	22101	36.12	18640	35.91

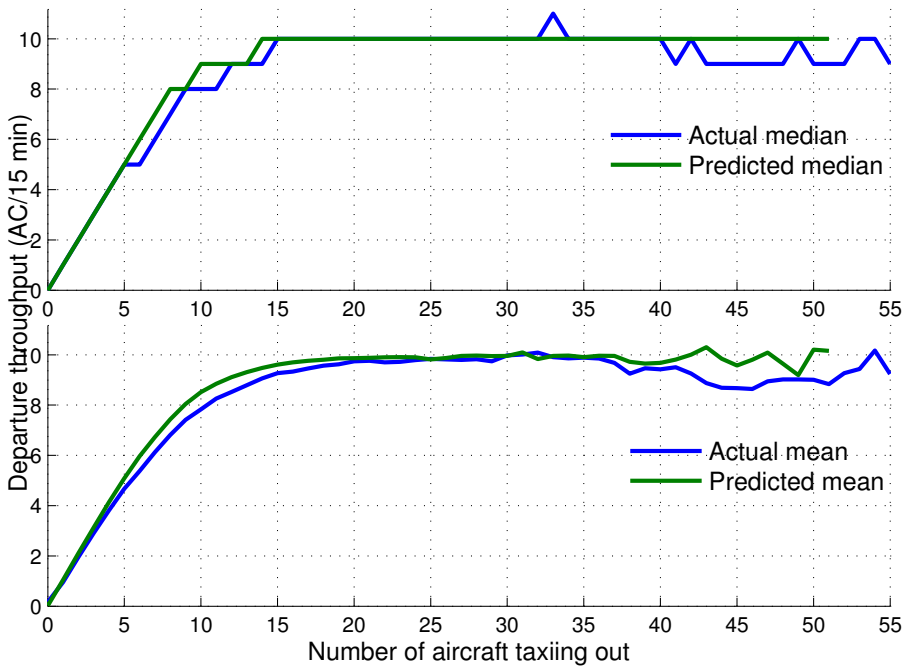


Figure 4-25: Actual and modeled throughput of all states N for EWR runway configuration 22L | 22R in year 2007: Mean (top); Median (bottom).

Variability of queuing delays

The predictions of the model can still be used for evaluating the delay uncertainty of a very busy schedule, such as the one of EWR in 2007. In Figure 4-26, we show the scatter plot of the predicted expected queuing delay of each flight (in min) and its predicted variance (in min^2) for all flights in configuration 22L | 22R in the years 2007 and 2011. Figure 4-26 shows the impact of the much higher pushback demand in the year 2007 on the variability of queuing delays, as predicted by the model.

Table 4.16: Prediction statistics for the congestion state and the throughput for EWR runway configuration 22L | 22R in year 2007.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-1.27	2.94	5.84	-2.78	5.09	8.36
Throughput (AC/15 min)	0.01	1.32	1.89	-0.27	1.57	2.17

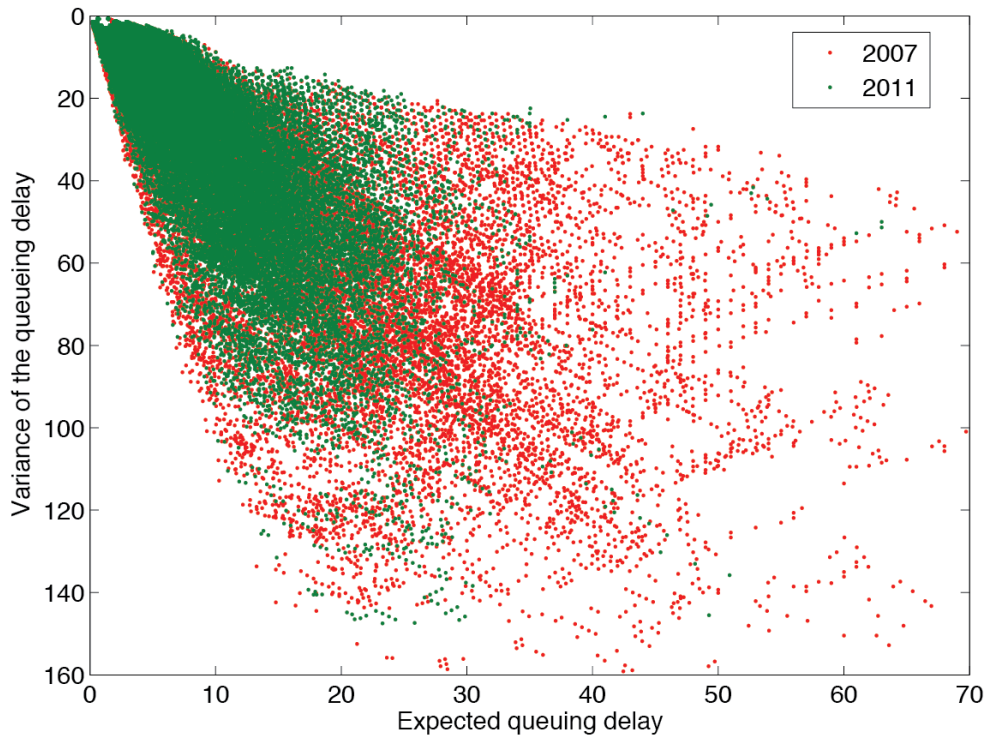


Figure 4-26: Predicted impact of high congestion on the variability of the queuing delays.

4.6.4 Predictions of taxi-out times of individual flights

In addition to aggregate comparisons, it is interesting to see how the model predicts individual taxi times. We compare the predicted taxi-out time for the flights for EWR configuration 22L | 22R in 2007, 2010 and 2011 to their recorded one. Figure 4-27 shows the cumulative distribution of the prediction error $E(l)$ for each flight l defined as

$$E(l) = \tau(l)^{pred} - \tau(l)^{obs} \tag{4.51}$$

In Figure 4-27, we observe that the error distribution is very similar for the years 2010 and 2011, and that for 88% of the flights, the taxi-out times are predicted within $\pm$ 10 minutes from the

recorded ones. The errors are much higher in 2007, and for only 73% of the flights are the taxi-out times are predicted within ± 10 minutes of the recorded ones. We also note that for approximately 2% of the flights (1000 flights) the taxi-out time is underestimated for 50 minutes or more in 2007. There are much fewer such flights in the years 2010 and 2011 (approximately 300). This change may have been related to the 3-hour tarmac delay rule that encourages airlines to cancel flights experiencing very long taxi-out times.

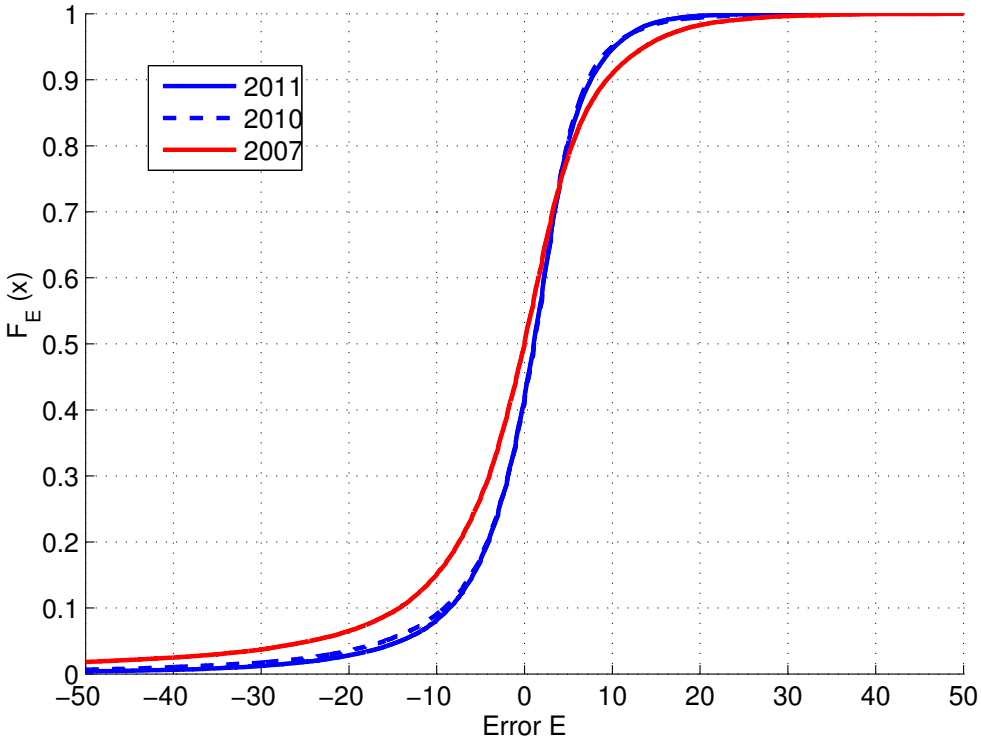


Figure 4-27: Taxi-out time prediction error for individual flights for EWR runway configuration 22L | 22R.

Table 4.17 lists the Mean Error (ME), Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) for these predictions. From Table 4.17, we observe that the model predicts taxi times of individual flights reasonably well in the years 2010 and 2011. The mean error is significant, which shows that there is some systematic underprediction in the predictions, which is by design, as explained earlier. In the following section, we investigate the reasons behind the weak prediction performance in 2007.

Table 4.17: Prediction statistics for individual taxi-out times for EWR runway configuration 22L | 22R in year 2007.

Year	ME (min)	MAE (min)	RMSE (min)
2011	-0.22	5.51	8.81
2010	-0.53	5.82	9.88
2007	-2.35	8.53	16.09

4.6.5 Airport performance in 2007

In Figure 4-28, we show the actual mean and median throughput measured in this runway configuration at EWR in the different years 2007, 2010 and 2011. It is evident that the throughput of the airport changes from years 2007 to 2011, especially at intermediate values of the number of aircraft taxiing out. Thus, although the average capacity of the airport stays at 10 AC/15 min across the three years, the expected throughput as a function of the number of aircraft taxiing out changes. We hypothesize that route availability, for which we do not recorded data for 2007, could drive this different behavior. Other possible reasons behind the difference across the three years include unimpeded taxi-out times, different local procedures, different magnitude of downstream constraints and traffic management initiatives, or regulations (for example the 3-hour tarmac delay rule).

4.6.6 Predictions for runway configuration 4R | 4L

We present the predictions for the second-most frequently used runway configuration at EWR in Appendix F. The methods used and the insights are similar to those discussed earlier in this chapter. We estimate the parameters using 2011 data, and predict operations in 2007 and 2010. Table 4.20 summarizes the main results for the individual flights taxi-out time predictions. The predictions are better than those for runway configuration 22L | 22R. For runway configuration 4R | 4L, the model predicts the average taxi-out times very accurately in all three years. 4R | 4L tends to not be used during convective weather, and thus the predictions are less impacted by the lack of route availability information in 2007 than those for runway configuration 22L | 22R. We also note that the individual flights taxi-out time prediction errors for 2010 (a testing dataset) are smaller than for 2011 (the training dataset).

Tables 4.19 and 4.20 summarize the actual mean taxi-out times for runway configurations 22L | 22R, 4R | 4L and those predicted by the model for the years 2007, 2010 and 2011. The model

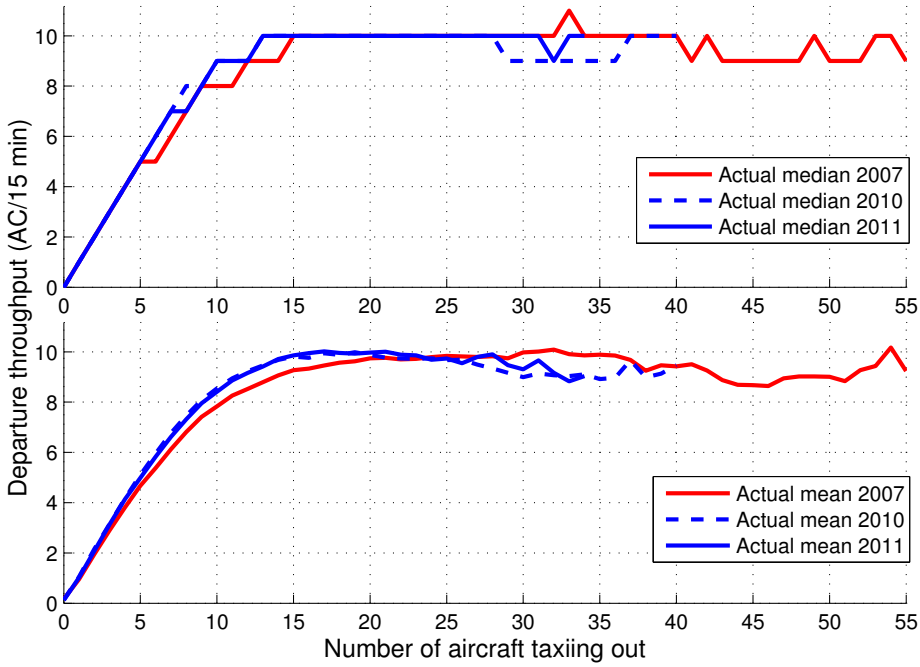


Figure 4-28: Actual throughput in all states N for EWR runway configuration 22L | 22R in years 2007, 2010 and 2011: Mean (top); Median (bottom).

Table 4.18: Prediction statistics for individual taxi-out times for EWR runway configuration 4R | 4L in years 2007, 2010 and 2011.

Year	Number of flights	Average taxi-out time (min)	ME	AME	RMSE
2011	37132	22.73	-0.50	5.58	8.67
2010	39785	22.86	0.17	5.48	8.17
2007	34378	29.55	-0.12	7.63	11.27

predicts the average airport performance reasonably well at both major runway configurations across different years.

4.7 Conclusions

In this chapter, we designed and validated a new analytical queuing model of the departure processes at airports that can be used for strategic planning and tactical predictions. For this, we developed a stochastic and dynamic queuing model of the departure runway(s), based on the transient analysis of $D(t)/E_k(t)/1$ queuing systems. A new method for estimating the unimpeded taxi times distribution, and a consistent method to estimate both the distribution of inter-departure

Table 4.19: Summarized prediction results for EWR runway configuration 22L | 22R in years 2007, 2010 and 2011.

Year	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Actual median taxi-out time (min)	Model median taxi-out time (min)
2011	20.41	20.19	18	18
2010	20.83	20.30	18	18
2007	25.83	23.48	21	19

Table 4.20: Summarized prediction results for EWR runway configuration 4R | 4L in years 2007, 2010 and 2011.

Year	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Actual median taxi-out time (min)	Model median taxi-out time (min)
2011	22.73	22.23	20	20
2010	22.86	23.03	20	21
2007	29.55	29.43	25	24

times and the throughput distribution were also proposed. The model was validated against Monte Carlo simulations and real data. It predicted taxi-out times, airport throughput and airport congestion level very accurately for the two major runway configurations of EWR in the years 2007, 2010 and 2011. In addition, we estimated the variance of the predicted queuing delays, and showed how these estimates can be used for evaluating operational uncertainty.

Chapter 5

Dynamic Control of Airport Departures

5.1 Introduction

In this chapter, we formulate the airport surface congestion management problem as a dynamic control problem, and present our findings both in terms of algorithmic development and field evaluation. Firstly, we develop a queuing model for the prediction of the departure throughput and derive two Pushback Rate Control (PRC) algorithms using dynamic programing and approximate dynamic programming (henceforth referred to as PRC_v2.0, and PRC_v2.1). Then, we describe the design of a Decision Support Tool that was used by air traffic controllers during the field testing of PRC_v2.1 at BOS in 2011. Subsequently, we describe the extensive evaluation of the congestion management scheme, the accuracy of the departure throughput prediction, and the implementation of PRC_v2.1. Finally, we describe an extension of these algorithms to a different airport environment, simulate its performance at PHL, and compare it to two other popular control mechanisms: N-Control and Slot-Control.

A large portion of the work presented in this chapter has appeared in conference publications [109, 110, 111].

5.2 Design of the control strategy

The objective of the control strategy is to minimize the amount of taxiing-out traffic, and thus taxi-out times, while maintaining runway utilization. In addition, it must be compatible with

current levels of information and automation in the airport tower, and capable of integration with current operational procedures, with minimal controller workload. Thus, the proposed strategy does not require Collaborative Decision Making, and does not assume the ability to plan and resequence departures. Its design has to address the uncertainties in the entire taxi-out process, from call-ready to takeoff.

For these reasons, the desired form of a congestion control strategy is one that periodically recommends a pushback (release) rate to air traffic controllers. The suggested pushback rate is updated at the beginning of each time-window, and is valid through that time period.

Careful monitoring of off-nominal events and constraints is also necessary for implementation at a particular site. In the case of BOS, of particular concern are gate conflicts (for example, an arriving aircraft is assigned the same gate as a departure that is being held), and the ability to meet controlled departure times (Expected Departure Clearance Times or EDCTs) and other constraints from Traffic Management Initiatives. In consultation with the BOS ATCT, it was decided that flights with EDCTs would be handled as usual and released First-Come-First-Served. Similarly, pushbacks would be expedited to allow arrivals to use the gate if needed. Finally, the analysis in Chapter 2 showed that, at BOS, prop departures do not interfere with jet departures and increase the total departure throughput. The main implication of this observation for the control strategy design at BOS is that props are exempt from the PRC.

Similarly, qualitative and quantitative observations show that there are differences in the departure process in the morning and evening periods [25]. In the morning, it is impacted by many EDCTs, whereas in the evening by heavy arrival pushes. For these reasons, we refine the data analysis and the departure throughput estimation for the evening times. The resultant saturation curve is shown in Figure 5-1.

5.2.1 Data sources

Figure 5-1 was determined using ASDE-X data, while the jet throughput analysis in Section 2.6 was performed with ASPM data. Pushback times in ASPM are based on the brake release times reported through the ACARS system, and are prone to error. About 40% of the flights departing from BOS do not automatically report these times [106]. While the ASDE-X data is more accurate than the ASPM data, it is still noisy, due to factors such as late transponder capture (the ASDE-X tracks only begin after the pilot has turned on the transponder, which may be before or after the actual pushback time), aborted takeoffs (which may have multiple departure times), flights cancelled

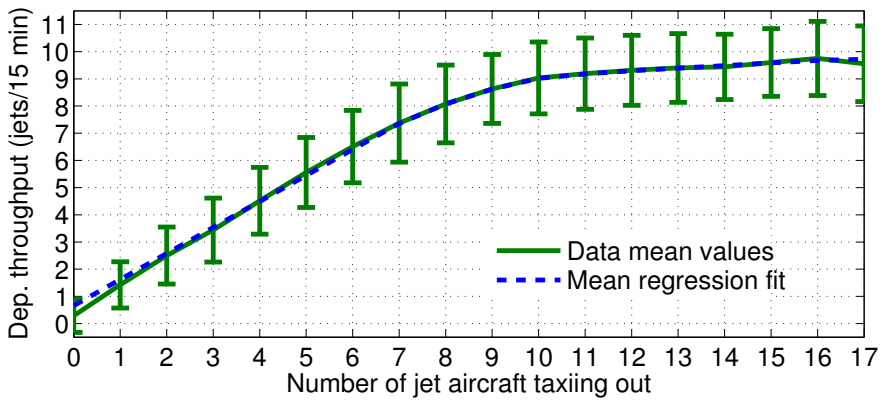


Figure 5-1: Regression of the departure throughput as a function of the number of aircraft taxiing out, for the 22L, 27 | 22L, 22R configuration at BOS during evenings, under VMC, using ASDE-X data.

after pushback, etc. A comparison of both ASDE-X and ASPM records with live observations made in the BOS tower on August 26, 2010 revealed that the average difference between the number of pushbacks per 15-minutes as recorded by ASDE-X and by visual means is 0.42, while it is -3.25 for ASPM and visual observations, showing that the ASPM records differ considerably from ASDE-X and live observations. The above comparison motivates the recalibration of airport performance curves and parameters using ASDE-X data in addition to ASPM data [110]. However, a comparison of ASDE-X records with live observations made in the tower of BOS during the evenings of July 7 – July 12, 2011 revealed that the average delay between the time that an aircraft was authorized to push back and the time at which it was captured in ASDE-X was 4 minutes. This delay was also shown to be random and not correlated with airline, aircraft type, or other information and introduces significant noise in the measurements of taxi-out times.

5.3 Departure process model

5.3.1 State variables

The beginning of each time-window is called *epoch*. At each epoch, we observe the state of the airport system, and recommend a pushback rate. For the purposes of control, the state is described by the following variables:

1. Visibility conditions and runway configuration $(MC; RC)$.
2. Number of jet aircraft traveling from the gates to the departure runway (G).
3. Number of jet aircraft in the departure queue (D).

4. Expected number of arrivals in the next 15 min. (A).
5. Number of props taxiing out (P).

All these variables are readily available in the current tower environment: G corresponds to the number of jet aircraft strips in the ground controller's rack, D is the number of jet aircraft strips local controller's rack, P can be determined visually from the same racks, and A can be determined from the Traffic Situation Display (TSD). The state of the departure process has to described using all or a subset of these variables. It is clear that a high-cardinality state space like the one proposed by Burgain [15], or assumed as input for the MILP models [75, 98, 115] cannot be applied in such an environment.

For developing the basic control algorithm PRC_v2.0, we use only two variables to describe the state of the airport given the current segment in use ($MC; RC$): At any time t, the state N_t of the departure process consists of the number of jet aircraft traveling from the gates to the departure queue (G_t) and the number of aircraft in the departure queue (D_t):

$$N_t = (G_t, D_t) \tag{5.1}$$

$$W_t = G_t + D_t \tag{5.2}$$

W_t is the total number of aircraft taxiing out, also known as the total work-in-process of the departure process.

5.3.2 Selection of time period

In general, the length of the time period, Δ, should be equal to the *lead time* of the system, that is, the delay between the application of the control input (setting an arrival rate at the runway server by controlling the pushback rate) and the time at which the runway sees that arrival rate. For the departure process, this time delay is given by the travel time from the gates to the departure queue. By choosing a time horizon that is approximately equal to the expected travel time from the gates to the departure queue, the flights released from the gate during a given time period are expected to reach the departure queue in the next time period.

In this way, we achieve meaningful relations among the variables (G, D) that describe the surface of the airport at each epoch, and the decision variable λ. Suppose, that at epoch τ, the system is at state (G_τ, D_τ) and the rate λ_τ is selected. Then, $G_{\tau+\Delta}$ correlates very well with λ_τ, and $D_{\tau+\Delta}$ with (G_τ, D_τ).

5.3.3 Pushback process

At the beginning of each time period, the decision maker chooses a pushback rate (arrival rate into the surface system), $\lambda \in \Lambda = [0, \lambda_{max}]$. λ is expressed as the number of pushbacks per Δ minutes. The time instances at which the pushback rate is updated are called *epochs*. In contrast to typical dynamic queuing control problems in which the decision maker sets the arrival rate into a facility, in our case, when setting a pushback rate at epoch τ, the decision maker authorizes λ aircraft to push back in that time period. In other words, λ pushbacks will occur in the time period $(\tau, \tau + \Delta]$ with probability 1 (w.p. 1). Furthermore, λ is an integer: $\lambda \in [0, 1, \ldots, \lambda_{max}]$.

5.3.4 Runway service process

The model treats the departure runways as a single server at which aircraft line up (queue) to await takeoff. The queuing system has finite queuing space C, which depends on the airport layout and operational procedures. At each airport, there is an upper bound on the number of aircraft that can queue up, which is the queuing space C of the queuing system. The runway service times are modeled as being Erlang distributed. The shape and rate $(k, k\mu)$ of the distribution are extracted from surveillance (ASDE-X) data, as explained in Section 4.4.3. The arrival times at the queuing system are modeled to be random and independent from each other. However, at each epoch, the total number or aircraft traveling from the gate to the departure queue is known (denoted R_τ). We assume that by the next epoch, all of them (R_τ) will have reached the runway server. We show later how this assumption can be relaxed. In summary, the arrival process at the runway is modeled as a non-stationary Poisson process, in which the rate is updated every Δ minutes, and the process is conditioned on the number of arrivals at the runway between two epochs.

This departure runway queuing system resembles a $M(t)/E_k/1$ system of queuing space C, with the additional constraint of R_τ arrivals during the $(\tau, \tau + \Delta]$ time interval. We denote it $(M(t)|R_\tau)/E_k/1$. Assuming that at epoch τ, R_τ aircraft are traveling to the departure runway, the probability density function g of the rth arrival at the departure runway at time t is:

$$\begin{aligned} g(r,t) &= \tfrac{R_\tau - (r-1)}{(\tau+\Delta)-t},\ t \in (\tau, \tau+\Delta],\ r = 0, 1, \ldots R_\tau \\ &= \tfrac{R_0 - (r-1)}{\Delta - t},\ \text{for } \tau = 0,\ t \in (0, \Delta],\ r = 0, \ldots R_0 \end{aligned} \tag{5.3}$$

To derive Equation (5.3), we consider $R_0 - (r-1)$ uniformly distributed random variables in the time interval $(t, \Delta]$. The probability that one of these lies in the interval $(t, t + dt]$ is $\big(R_0 - (r -$

$1)\big)dt/(\Delta - t)$.

The state of the queuing system at time t is denoted by

$$S_t = (R_t, Q_t) \tag{5.4}$$

where R_t is the number of aircraft that were traveling to the departure runway at the start of that epoch but have not reached the departure queue yet, and $Q_t \in \{0, 1, \ldots, kC\}$ is the state of the embedded chain of the semi-Markov process. An example of the chain for $k = 2$ and $C = 4$ is shown in Figure 5-2.

A service completion of an Erlang process with shape k and rate $k\mu$ is represented with k stages of exponentially distributed random variables with rate $k\mu$. We call each such stage *stage of work*, as explained in Section 4.4.1. Each state of the Markov chain (r, q) denotes that there are r aircraft that have been traveling to the runway since the start of that epoch, and there are q *stages of work* to be completed at the departure runway server, i.e., there are $\min(1, q)$ aircraft in service and $\max(\lceil (q-k)/k \rceil, 0)$ aircraft in the departure queue.

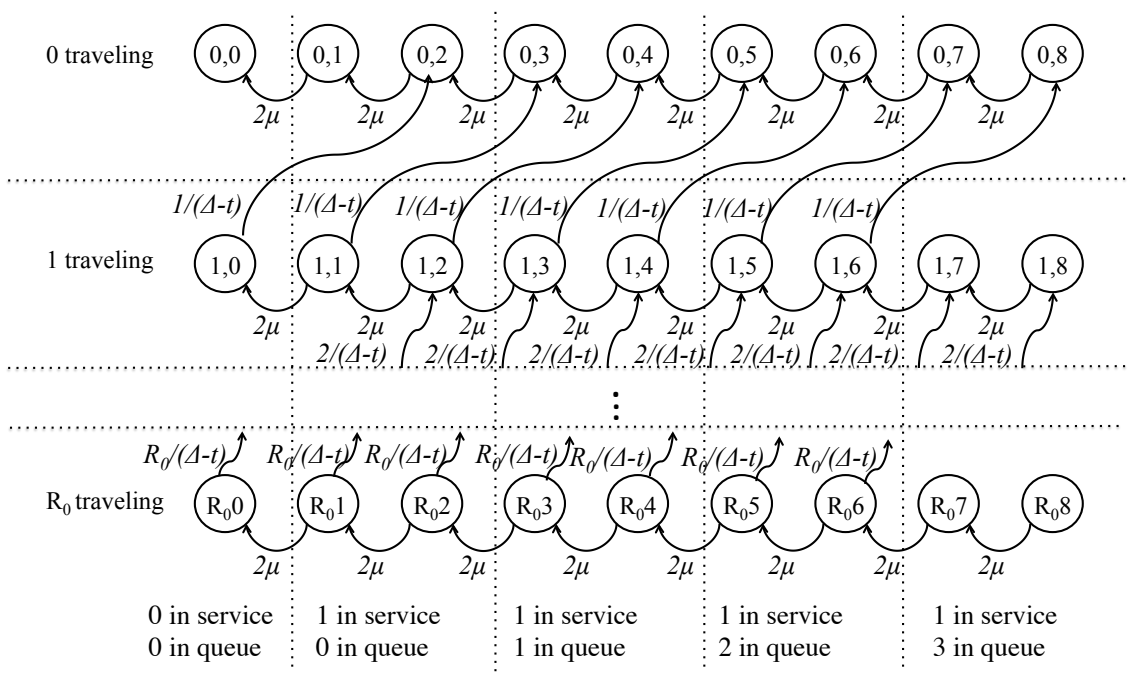


Figure 5-2: State transition diagram for an $(M(t)|R_0)/E_2/1$ system with queuing space of 4 customers in the system.

At epoch 0, the Markov chain is in state (R_0, Q_0). In Figure 5-2, the chain is in the bottom

level (R_0 aircraft traveling to the departure runway) with Q_0 stages of work to be completed. By the end of the time interval Δ, all of R_0 aircraft will have reached the departure queue, and the Markov chain will be at the top level (0 aircraft traveling). Let $P_{r,q}(t)$ denote the probability that the queuing system is in state (r, q) at time t, where $0 < t \leq \Delta$. The state probabilities $P_{0,0}(\Delta), P_{0,1}(\Delta), \cdots P_{0,kC}(\Delta)$ describe fully the state of the queuing system at the end of the time interval Δ. They are calculated by deriving the first-order differential equations (*Chapman-Kolmogorov equations*) that describe the evolution over the time $(0, \Delta]$, given R_0 arrivals in this interval: For $0 < t \leq \Delta$, and $1 \leq r < R_0$:

$$\frac{dP_{0,0}}{dt} = k\mu P_{0,1} \tag{5.5}$$

$$\frac{dP_{0,q}}{dt} = k\mu P_{0,q+1} - k\mu P_{0,q}, \quad 1 \leq q < k \tag{5.6}$$

$$\frac{dP_{0,q}}{dt} = k\mu P_{0,q+1} + \frac{1}{\Delta - t} P_{1,q-k} - k\mu P_{0,q}, \quad k \leq q < kC \tag{5.7}$$

$$\frac{dP_{0,kC}}{dt} = \frac{1}{\Delta - t} P_{1,k(C-1)} - k\mu P_{0,kC} \tag{5.8}$$

$$\frac{dP_{r,0}}{dt} = k\mu P_{r,1} - \frac{r}{\Delta - t} P_{r,0} \tag{5.9}$$

$$\frac{dP_{r,q}}{dt} = k\mu P_{r,q+1} - k\mu P_{r,q} - \frac{r}{\Delta - t} P_{r,q}, \quad 1 \leq q < k \tag{5.10}$$

$$\frac{dP_{r,q}}{dt} = k\mu P_{r,q+1} + \frac{r+1}{\Delta - t} P_{r+1,q-k} - \frac{r}{\Delta - t} P_{r,q} - k\mu P_{r,q}, \quad k \leq q \leq k(C-1) \tag{5.11}$$

$$\frac{dP_{r,q}}{dt} = k\mu P_{r,q+1} + \frac{r+1}{\Delta - t} P_{r+1,q-k} - k\mu P_{r,q}, \quad k(C-1) < q < kC$$

$$\frac{dP_{r,kC}}{dt} = \frac{r+1}{\Delta - t} P_{r+1,k(C-1))} - k\mu P_{r,kC} \tag{5.12}$$

$$\frac{dP_{R_0,0}}{dt} = k\mu P_{R_0,1} - \frac{R_0}{\Delta - t} P_{R_0,0} \tag{5.13}$$

$$\frac{dP_{R_0,q}}{dt} = k\mu P_{R_0,q+1} - \left(\frac{R_0}{\Delta - t} - k\mu\right) P_{R_0,q}, \; 1 \leq q \leq k(C-1) \tag{5.14}$$

$$\frac{dP_{R_0,q}}{dt} = k\mu P_{R_0,q+1} - k\mu P_{R_0,q}, \; k(C-1) < q < kC \tag{5.15}$$

$$\frac{dP_{R_0,kC}}{dt} = -k\mu P_{R_0,kC} \tag{5.16}$$

Solving Equations (5.5)-(5.16) numerically for time $t = \Delta$ with initial value (R_0, Q_0), we obtain the state probabilities $P_{0,0}(\Delta), P_{0,1}(\Delta), ...P_{0,kC}(\Delta)$. The state of the queuing system at time Δ, Q_Δ, is a probabilistic function f of the initial value (R_0, Q_0), and the probabilities $p_{q(i)}$ of each

state i are the calculated probabilities $P_{0,i}(\Delta)$:

$$Q_\Delta = f(R_0, Q_0) \tag{5.17}$$

$$\text{with} \quad p_{q(i)}(R_0, Q_0) = P_{0,i}(\Delta) \quad \text{for } 0 \leq i \leq kC \tag{5.18}$$

$$\implies \quad \mathbf{p}_q(R_0, Q_0) = \mathbf{P}_0(\Delta) \tag{5.19}$$

where $\mathbf{P}_0(\Delta) = [P_{0,0}(\Delta), P_{0,1}(\Delta), ...P_{0,kC}(\Delta)]'$.

5.3.5 System dynamics

Suppose, at epoch τ, that R_τ aircraft are traveling to the departure runway, Q_τ stages of work are left to be completed in the queue, and the decision maker selects a pushback rate λ_τ. At $\tau + \Delta$, R_τ aircraft will have reached the departure queue, λ_τ aircraft will be traveling, and $Q_{\tau+\Delta} = f(R_\tau, Q_\tau)$ stages of work will remain to be completed. The queuing system therefore evolves according to the following equation:

$$(R_{\tau+\Delta}, Q_{\tau+\Delta}) = (\lambda_\tau, f(R_\tau, Q_\tau)) \tag{5.20}$$

The probabilities $\Pr_{(r,q)\to(i,j)}(\lambda)$ that the chain is in state (i, j) at the next epoch $\tau + \Delta$ given it is in state (r, q) at the epoch τ and the pushback rate λ is chosen are:

$$\Pr_{(r,q)\to(i,j)}(\lambda) = \begin{cases} p_{q(j)}(r, q) & \text{if } i = \lambda \\ 0 & \text{otherwise} \end{cases} \tag{5.21}$$

The state S of the queuing system maps to the state N of the departure process as follows:

$$N_t = \begin{cases} \big(\lambda_{t-\Delta}, \max(\lceil (Q_t - k)/k \rceil, 0)\big), & t \in \{0, \Delta, \dots\} \\ \big(V_t + R_t, \max(\lceil (Q_t - k)/k \rceil, 0)\big), & \text{otherwise} \end{cases} \tag{5.22}$$

where V_t is the number of aircraft that pushed back between the start of the epoch within which t lies, and the time t. We note that by sampling the system every Δ time intervals, we decouple the departure process into two processes that are independent of each other within each time period, namely, the pushback process and the runway service process.

5.3.6 Choice of cost function

The control strategy sets the arrival rate of aircraft to the queuing system, that is the pushback rate, to balance two objectives, namely, to minimize the expected departure queue length and to maximize the runway utilization. These requirements are captured in a cost function, $c(q)$ for a state (r, q) of the queuing system. This cost is a combination of the queuing cost and the cost of non-utilization of the runway. The runway is unutilized when $q = 0$. If $q \in \{1, 2, \ldots k\}$, both the queuing and non-utilization costs are zero. For all higher states, $q > k$, there is a queuing cost $c(q)$, which is usually assumed to be a monotonically non-decreasing function of q with increasing marginal costs [78, 80]. In this case, it scales quadratically with the state of the queue, because the expected system delay scales as a quadratic function of queuing state ($[D \cdot (D+1)/2]/\mu$). A candidate cost function with these properties is:

$$c(q) = \begin{cases} H, & q = 0 \\ (\lceil (q-k)/k \rceil)^2 & q = 1, \ldots, kC \end{cases} \tag{5.23}$$

where H is the cost of a loss of runway utilization.

We solve Equations (5.5)-(5.16) numerically to calculate

$$\mathbf{p}_q(R_0, Q_0, t) = \big[\sum_{r=0}^{R_0} P_{r,0}(t), \sum_{r=0}^{R_0} P_{r,1}(t), \ldots, \sum_{r=0}^{R_0} P_{r,kC}(t)\big]' \tag{5.24}$$

at time t. Numerical experiments showed that sampling every 6 sec (*i.e.* 10 times a minute) is sufficiently accurate for calculating the expected cost of each state, $\bar{c}$, over the time interval Δ:

$$\bar{c}(R_0, Q_0) = \sum_{i=0}^{10\Delta-1} \frac{1}{10} \mathbf{p}_q(R, Q, i/10) \cdot \mathbf{c} \tag{5.25}$$

5.4 Dynamic programing formulation

The optimal costs, $J^*(r, q)$, at each state, (r, q), for the infinite horizon problem with discount factor α are given by Bellman's equation:

$$\begin{aligned} J^*(r,q) &= \min_{\lambda \in \Lambda} \{\bar{c}(r,q) + \alpha \sum_{j=0}^{kC} \Pr_{(r,q)\to(\lambda,j)} J^*(\lambda, j)\} \\ \Longrightarrow J^*(r,q) &= \min_{\lambda \in \Lambda} \{\bar{c}(r,q) + \alpha \mathbf{p}_q(r,q) \cdot \mathbf{J}^*(\lambda) \end{aligned} \tag{5.26}$$

where $\mathbf{J}^*(\lambda) = [J^*(\lambda, 0), J^*(\lambda, 1), \ldots, J^*(\lambda, kC)]'$ for $r \in \{0, 1, \ldots, \lambda_{max}\}$ and $q \in \{0, 1, \ldots, kC\}$.

We now relax the assumption of Equation (5.20) that R_τ aircraft traveling at epoch τ will reach the queue during the time interval $(\tau, \tau + \Delta]$ and a pushback rate (λ_τ) set at epoch τ will arrive at the runway at $t > \tau + \Delta$ w.p. 1, as follows. For each value of λ_τ and R_τ, i out of the λ_τ aircraft reach the runway during the time interval $(\tau, \tau + \Delta]$, with probability β_i. Similarly, i out of the R_τ aircraft reach the runway at $t > \tau + \Delta$ with probability γ_i. Finally, R_τ aircraft reach the runway during the time interval $(\tau, \tau + \Delta]$, and λ_τ aircraft at $t > \tau + \Delta$, with probability $1 - \sum \beta_i - \sum \gamma_i$.

Equation (5.20) becomes:

$$(R_{\tau+\Delta}, Q_{\tau+\Delta}) = \begin{cases} \big(\lambda_\tau, f(R_\tau, Q_\tau)\big), & \text{w.p. } 1 - \sum \beta_i - \sum \gamma_i \\ \big(\lambda_\tau - i, f(R_\tau + i, D_\tau)\big), \text{w.p. } \beta_i, i = 1, \ldots, \lambda_\tau \\ \big(\lambda_\tau + i, f(R_\tau - i, D_\tau)\big), \text{w.p. } \gamma_i, i = 1, \ldots, R_\tau \end{cases} \tag{5.27}$$

We note that Equation (5.27) maintains the Markov property. For these system dynamics, the Bellman equation for the infinite horizon problem with discount factor α is:

$$\begin{aligned} J^*(r, q) = \min_{\lambda \in \Lambda} \Big\{ & (1 - \sum \beta_i - \sum \gamma_i)[\bar{c}(r, q) + \alpha \mathbf{p}_q(r, q) \cdot \mathbf{J}^*(\lambda)] \\ & + \sum \beta_i [\bar{c}(r + i, q) + \alpha \mathbf{p}_q(r + i, q) \cdot \mathbf{J}^*(\lambda - i)] \\ & + \sum \gamma_i [\bar{c}(r - i, q) + \alpha \mathbf{p}_q(r - i, q) \cdot \mathbf{J}^*(\lambda + i)] \Big\} \end{aligned} \tag{5.28}$$

Equation (5.28) illustrates the tradeoffs involved with the choice of appropriate time period, Δ. If the time period is large, less frequent updates of the optimal policy are necessary, which makes implementation easier. On the other hand, it is then necessary to predict runway performance and maintain runway utilization over a longer period of time, which may lead to less effective congestion control. If the time period is significantly shorter than the lead time, finer control is theoretically possible. Thus, a smaller inventory will be necessary at the departure queue to maintain runway utilization. On the other hand, a larger number of aircraft traveling will be necessary at each epoch, because only a fraction of them will have arrived at the runway by the next epoch (that is, large values of γ_i). In other words, the state at epoch τ, (R_τ, Q_τ), will correlate poorly with the state of the queue at the next epoch, $Q_{\tau+\Delta}$. Thus, it is not clear if the policy will be more effective at controlling congestion. In addition, more frequent updates of the optimal policy will be necessary, which will increase the workload of air traffic controllers.

Finally, we note that this problem satisfies the property of weak accessibility: Suppose that at the beginning of epoch 0, the embedded chain is at state (r_0, q_0). At the beginning of the next epoch the chain will be at any of the states $(\lambda_0, 0), (\lambda_0, 1), \ldots (\lambda_0, \min(r_0 + q_0, kC))$ with non-zero probability. Suppose that the following control law is applied: For all (r_0, q_0), $\lambda_0 = \lambda_{max}$, where $\lambda_{max} > \mu$. Then, the queuing system will reach the state (λ_{max}, kC) within a finite number of epochs with nonzero probability. Also, at the next epoch, the state will be in any of the states $(\lambda_{max}, 0), (\lambda_{max}, 1), \ldots (\lambda_{max}, kC))$ with nonzero probability. As before, from any of these states, the chain will reach the state (λ_{max}, kC) within a finite number of epochs with nonzero probability. Therefore, the state (λ_{max}, kC) is recurrent under this control law, and weak accessibility is satisfied.

Using a discount factor as in Equation (5.28) may not be appropriate, since the cost of an unutilized runway remains constant in time. An alternate formulation is to determine the policies that minimize the average optimal cost per stage, c^*:

$$\begin{aligned} c^* + h^*(r,q) \quad = \min_{\lambda \in \Lambda} \Big\{ \quad & (1 - \sum \beta_i - \sum \gamma_i)[\bar{c}(r,q) + \mathbf{p}_q(r,q) \cdot \mathbf{h}^*(\lambda)] \\ & + \sum \beta_i [\bar{c}(r+i,q) + \mathbf{p}_q(r+i,q) \cdot \mathbf{h}^*(\lambda - i)] \\ & + \sum \gamma_i [\bar{c}(r-i,q) + \mathbf{p}_q(r-i,q) \cdot \mathbf{h}^*(\lambda + i)] \Big\} \end{aligned} \tag{5.29}$$

5.5 Application of PRC at BOS

This section describes the application of PRC_v2.0, as given by Equation (5.29), to the departure process at BOS. We focus on runway configuration 22L, 27 | 22L, 22R in VMC during the evening departure push. The control strategy is applied only to jet aircraft at BOS, for reasons explained in Section 5.2. We also refined PRC_v2.0 for runway configuration 4L, 4R | 4L, 4R, 9 in VMC during the evening departure push and 4R | 4R, 9 in IMC during the evening departure push.

5.5.1 Selection of time period

The average unimpeded taxi-out time at BOS is 12.6 minutes under VMC [107]. There is an added delay due to taxiway congestion, which is proportional to the number of aircraft traveling to the runway[70, 107]. For non-excessive traffic levels, the additional average delay in the case of the BOS airport is 1-2 minutes. This makes 15 minutes a suitable choice of time-window for BOS. Furthermore, we assume that $\beta_i = \gamma_i = 0$ for all i because of lack of accurate measurements, as

explained in 5.2.1. Equation (5.29) then becomes:

$$c^* + h^*(r,q) = \min_{\lambda \in \Lambda} \{ (\bar{c}(r,q) + \mathbf{p_q}(r,q) \cdot \mathbf{h}^*(\lambda) \} \tag{5.30}$$

5.5.2 Estimation of runway service process parameters

We are interested in estimating the parameters of the runway service process of the BOS airport during peak evening times. For this reason, we perform the analysis outlined in Chapter 2 with ASDE-X data from November 2010-June 2011, and isolate 15-minute intervals in the filtered dataset in saturation. We obtain 1726 measurements of the runway throughput (AC/15 min) that provide an empirical distribution of the departure capacity in the evening times. Figure 5-3 shows the resulting empirical distribution f_{rw} in black.

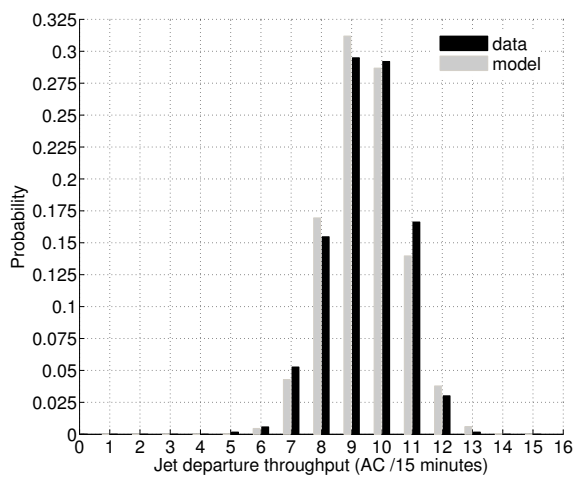


Figure 5-3: Empirical (f_{rw}) and modeled (f_{rm}) probability distributions of the departure capacity of runway configuration 22L, 27|22L,22R under visual meteorological conditions during evening times.

An Erlang distribution, f_{rw}, is fitted using the approximate Method of Moments outlined in Section 4.4.3. For the empirical distribution of Figure 5-3, we obtain the parameters of Erlang distribution (5, 3.25). The mean service time is $5/3.25 = 1.54$ min. The variance of the service time is 0.47 min^2. The corresponding distribution, f_{rm}, of the number of departures in 15 min is depicted in Figure 5-3 in grey. The empirical and modeled distributions are similar, as can also be seen in Table 5.1.

Table 5.1: First two moments of the distributions f_{rw} and f_{rm}.

	f_{rw}	f_{rm}
μ_1(AC/min)	9.74	9.74
μ_2(AC/min)^2	96.80	96.92

5.5.3 Maximum pushback rate and cost function

The set of permissible policies is defined as $0, 1, \ldots, \lambda_{max}$. At BOS, as in most airports, there is a natural threshold for the maximum admissible rate of arrivals into the departure process (pushbacks). At BOS, λ_{max} is calculated to be 15 AC/15 min, that is, $\Lambda = \{0, 1, \ldots, 15\}$. The space of the queuing system (C) is estimated to be 30, and the cost of underutilizing the runway, c(0), is chosen to be equal to the cost of a queue of 25 departures. $c(0)$ is chosen to reflect the fact that at BOS, a very long queue can lead to surface gridlock, and consequently, non-utilization of the runway.

5.5.4 Calculation of optimal policies

Given the service time distribution $(k, k\mu)$, the time period Δ, the queuing space C, the set Λ and the costs c, Equation (5.30) can be solved to obtain the optimal pushback policies. The efficient solution of Equation (5.30) is possible using the policy iteration method with a suitable choice of initial policy. In selecting initial policies, we use the insights that (1) For given q, the pushback policy is expected to be a non-decreasing function of r; (2) For given r, the pushback policy is expected to be a non-decreasing function of q; (3) The pushback policy is expected to target for a specific level of inventory (number of aircraft in the queue). (4) The pushback rates take values between 0 and λ_{max}. We used a target inventory, $b_f = 5$ aircraft in the queue. For each state (r, q), the initial policy $\lambda_0(r, q)$ is calculated as:

$$\lceil \min(\max(\mu + b_f - \max(r + \max(\lceil (q-k)/k \rceil, 0) - \mu, 0), 0), \lambda_{max}) \rceil$$

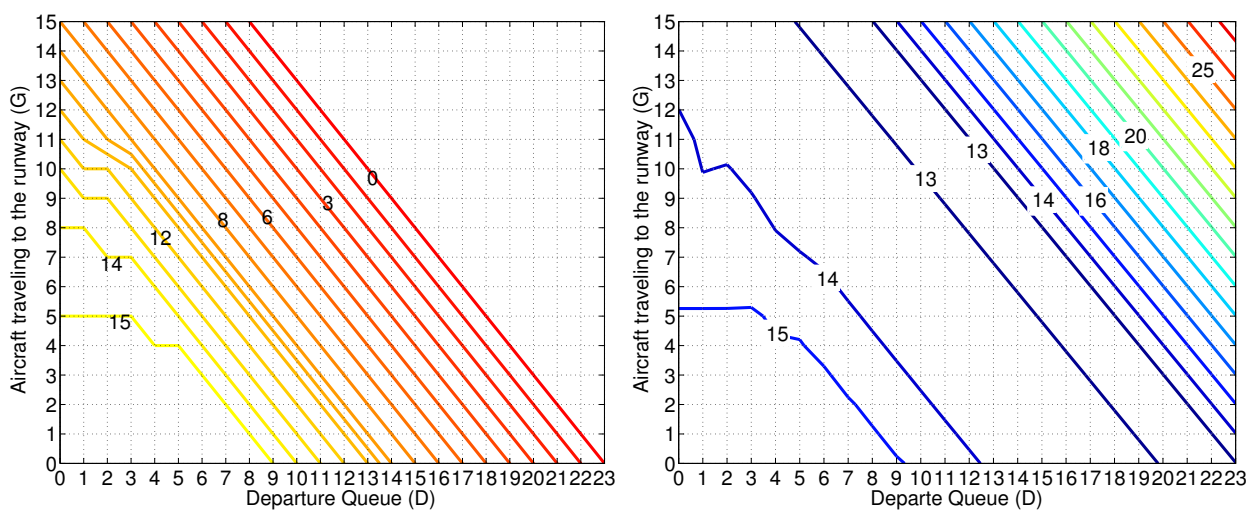
The policy iteration algorithm converges in fewer than 10 iterations. The optimal policies λ^* are a function of the state of the embedded chain (r, q), which is not observable. However, each state of the chain is mapped to an observed state of the process, N (Equation 5.22). For $0 \leq T \leq \lambda_{max}$,

the optimal pushback rate is approximated by:

$$\bar{\lambda}(G,0) = \lfloor \frac{\sum_{j=0}^{k} \lambda^*(G,j)}{k+1} + 0.5 \rfloor \tag{5.31}$$

$$\bar{\lambda}(G,D) = \lfloor \frac{\sum_{j=Dk+1}^{(D+1)k} \lambda^*(G,j)}{k} + 0.5 \rfloor \quad \text{for } 1 \leq D < C \tag{5.32}$$

Figure 5-4a shows the contours of the optimal pushback policy $\bar{\lambda}$ as a function of the number of aircraft in the departure queue (D) and the number of aircraft traveling to the runway (G). As expected, the optimal pushback rates decrease for increasing D and G. A different way to characterize the optimal policies is to plot the expected work-in-process at the next epoch, $\bar{W}_{\tau+\Delta} = G_{\tau+\Delta} + \bar{D}_{\tau+\Delta} = \bar{\lambda}_\tau + \bar{D}_{\tau+\Delta}$ (5.20) as a function of the current state (G_τ, D_τ), as shown in Figure 5-4b. When $W_\tau \geq 23$, the optimal pushback rate is 0, but it is not sufficient to reduce $\bar{W}_{\tau+\Delta}$ to 13. We also note that when $W_\tau \leq 13$, the optimal pushback policy increases $\bar{W}_{\tau+\Delta}$ to values higher than 13.



(a) Optimal pushback policy $\bar{\lambda}_\tau$ as a function of the number of aircraft in the departure queue (D_τ) and the number of aircraft traveling to the runway (G_τ).

(b) Expected work-in-process at the next epoch, $\bar{W}_{\tau+\Delta}$, as a function of the number of aircraft in the departure queue (D_τ) and the number of aircraft traveling to the runway (G_τ).

Figure 5-4: Optimal pushback policy and expected work-in-process as a function of the current state.

Figure 5-4b suggests that the algorithm aims at controlling the process to a desired value of W_τ. The expected work in process $\bar{W}_{\tau+\Delta}$ consists of the expected queue length at $\tau+\Delta$, $\bar{D}_{\tau+\Delta}$, and the pushback rate $\bar{\lambda}_\tau$ set at at time τ (Equation 5.22). Comparing Figures 5-4a and 5-4b, we observe

that each $\bar{\lambda}_\tau$ is associated with one value of $\bar{W}_{\tau+\Delta}$, or $\bar{D}_{\tau+\Delta}$. For example, when $\bar{W}_{\tau+\Delta} = 0$, $\bar{\lambda}_\tau$=15, and when $\bar{W}_{\tau+\Delta} = 13$, $\bar{\lambda}_\tau = 13 - \bar{D}_{\tau+\Delta}$. This implies that the optimal pushback policy at time τ, is a function of the expected queue length at time $\tau + \Delta$.

Figure 5-5 shows the scatterplot between the optimal pushback rate $\bar{\lambda}_\tau(G_\tau, D_\tau)$ and the expected $\bar{D}_{\tau+\Delta}(G_\tau, D_\tau)$, for all $0 \leq G \leq \lambda_{max}$ and $0 \leq D \leq C$, along with a fitted convex non-increasing function that minimizes absolute deviations from the calculated points. The equivalent PRC_v1.0 strategy [110], which aims at keeping $W_{\tau+\Delta}$ always at 13 irrespective of the state N_τ, is also shown. For the most part, the two strategies are the same after rounding to the closest integer. However, when the expected queue length at $\tau + \Delta$ is less than 4, the optimal pushback policy increases $W_{\tau+\Delta}$ to 14 or 15. In this region, the departure throughput can be increased with a high pushback rate at a very low congestion cost. Figure 5-5 also shows the benefit of the PRC_v1.0 strategy [110]. By simply aiming at a target $W_{\tau+\Delta}$ at the next epoch, the strategy is suboptimal only when the expected value of $W_{\tau+\Delta}$ is 1, 2 or 3. However, these are instances of high risk of runway non-utilization, and PRC_v2.0 accounts better for this risk.

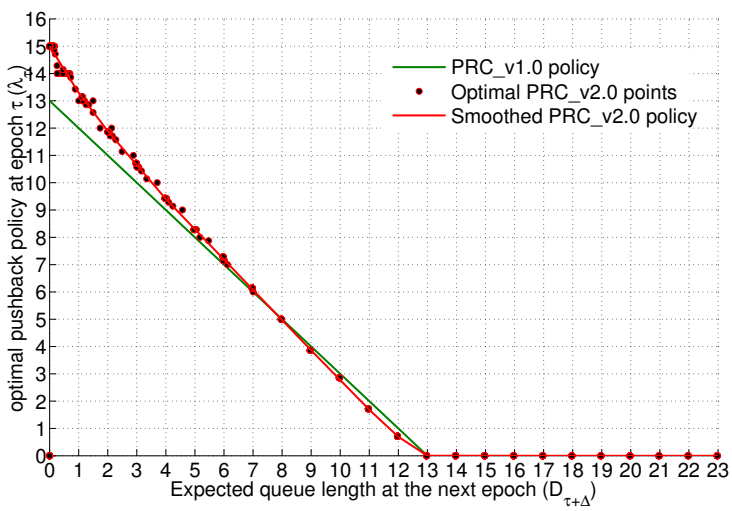


Figure 5-5: Optimal pushback policy $\bar{\lambda}_\tau$ as a function of the expected *queue* $\bar{D}_{\tau+\Delta}$ at the next epoch $(\tau + \Delta)$.

To illustrate how the control algorithm would work in conjunction with the system dynamics described in Equation (5.20), we consider a sample path of the certainty equivalent system: At the first epoch ($t = 0$), the state is $(0, 0)$, that is, there are no aircraft on the ground. At the next epoch ($t = 15$), the expected queue will be zero, and the curve of Figure 5-5 recommends that 15 aircraft pushback in the next 15 minutes (or a pushback rate of 1/ minute). Thus, $S_{15} = (15, 0)$. Solving the Chapman-Kolmogorov equations numerically for the queuing model $(M(t)|R_\tau)/E_k/1$, we find that

at the third epoch ($t = 30$), the expected queue length is 5. As a result, Figure 5-5 recommends a pushback rate of (8/15 minutes), so $S_{30} = (8, 5)$. Similarly, $S_{45} = (11, 3)$, $S_{60} = (9, 4)$, $S_{75} = (11, 3)$, etc. Therefore, after two cycles, the system stabilizes at a traffic level of 13-14 aircraft, and the queue at 3-4 aircraft. We also note that the target queue length at each epoch is at least 3. Finally, since the pushback rate is bounded at 15 AC/15 min, the traffic level can reach at most 24 aircraft: This happens in the extreme case in which the state is $(0, 10)$, which implies $\bar{\lambda} = 14$, and no aircraft manages to takeoff. If this happens, due to an unpredicted runway closure for example, the next state is $(14, 10)$ and the pushback rate is set to 0, as can be seen from Figure 5-13.

5.5.5 Conditional throughput forecasts

Parameters such as the fleet mix and the expected number of landings in the next time window $(\tau, \tau + \Delta]$ can provide a conditional forecast for the runway service time distribution as discussed in Section 4.4.3. These parameters explain some of the variance of the departure throughput and provide a better estimate of the expected departure capacity. For example, for runway configuration 22L,27 | 22L,22R at evening times under visual meteorological conditions, the departure throughput given arrival throughput (landings) and props demand can be estimated from the regression tree of Figure 5-6. This regression tree is validated using 10-fold cross validation.

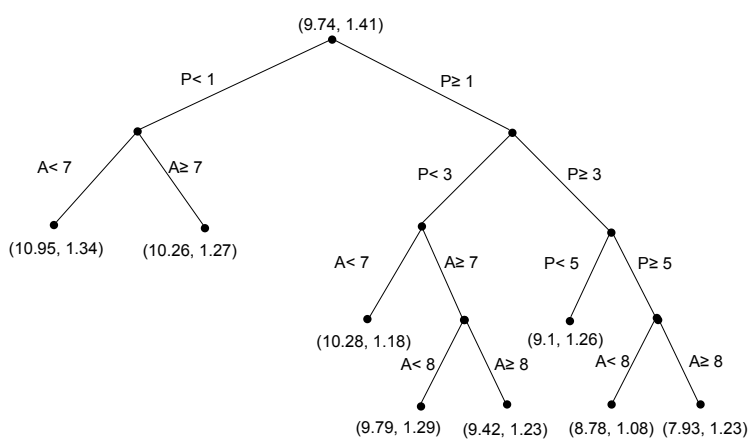


Figure 5-6: Jet departure throughput prediction (mean, standard deviation) given expected number of Arrivals in the next 15 minutes and number of props taxiing out.

These conditional forecasts are incorporated into the algorithm as follows:

- At epoch τ, calculate the conditional throughput distribution estimate for the time window $(\tau, \tau + \Delta]$ using the expected number of Arrivals (A) and the number of props taxiing out

(P) from the regression tree.

- Calculate the expected takeoff rate in the time window $(\tau, \tau+\Delta]$ and queue length at $\tau+\Delta$ using the queuing model $(M(t)|R_\tau)/E_k/1$ with parameters fitted to the conditional throughput forecast distribution

- Use the PRC_v2.0 curve (such as the one of Figure 5-5) to calculate the optimal pushback policy for this expected queue length.

This is a heuristic modification of PRC_v2.0 in the spirit of roll-out algorithms [11] to incorporate the conditional forecast. We call this control protocol PRC_v2.1. The intuition behind the derivation of PRC_v2.1 is that we use the conditional forecasts only for updating our belief for the expected queue. We do not take advantage of the fact that the reduction in the variance of the capacity distribution could imply a more aggressive control policy, that is a shorter target queue. Given that the conditional forecast has reduced variance compared to the unconditional one, we conjecture that PRC_v2.1 is more optimal than PRC_v2.0. We denote the optimal policies of PRC_v2.1 as $\check{\lambda}$. The optimal policy $\check{\lambda}_\tau$ is a function of the departure queue, the number of aircraft traveling to the runway , the number of props taxiing out at epoch τ and the expected number of landings in the time window $(\tau, \tau + \Delta]$. This heuristic is chosen because of its simplicity and intuitiveness. An alternative would be to augment the state and include the throughput forecast as a state variable.

We also note that another approach to predicting the takeoff rate would be a data-driven approach. In this approach, we would predict the jet departure throughput given the four explanatory variables G, D, A, P from statistical models derived from historical data. The simplest such model would be simply using the fitted throughput curve of Figure 5-1. This curve yields the expected jets throughput in the time window $(\tau, \tau + \Delta]$ given that there are currently $(G + D)$ jets taxiing out. We also develop a more sophisticated version of this model that considered all four explanatory variables using regression trees. A significant issue with these models is the delayed capture of the aircraft in ASDE-X. As explained in Section 5.2.1, aircraft appear in ASDE-X on average 4 minutes after they pushed back. Thus, when ASDE-X shows x aircraft taxiing, the actual number could be significantly higher. We discuss the performance of these models in Section 5.7.3.

5.5.6 Rounding of optimal policies

As explained in Section 5.2, it is required for the optimal policy to be given in form of a rate. Thus, the optimal pushback policy for each 15-minute time-period is rounded to one of the following rates:

0 AC/15 min, 1 AC/5 min, 1 AC/3 min, 2AC/5 min, 1 AC/2 min, 3AC/5 min, 2 AC/3 min, 4 AC/5 min, 1 AC/min.

5.6 Design of a Decision Support Tool

The next step of our research is the investigation of the downstream deployment potential of PRC algorithms. To this end, we develop an application that uses the necessary inputs to automatically determine the suggested rate. The design and development was joint work with M. Sandberg and is described in detail here [111]. The device used is a tablet computer, the 7-inch Samsung Galaxy Tab ™, which has the advantages of being portable and compact. In addition, the Android operating system offers a convenient application development environment. Two tablet computers are used for the implementation of the strategy; the rate control transmitter and the rate control receiver. The rate control transmitter is used to input the data, and the rate control receiver to display the recommended rate to the Boston Gate (BG) controller, who is responsible for authorizing aircraft to monitor ground control for their pushback. The two devices communicate with each other using a Bluetooth wireless link (Figure 5-7).

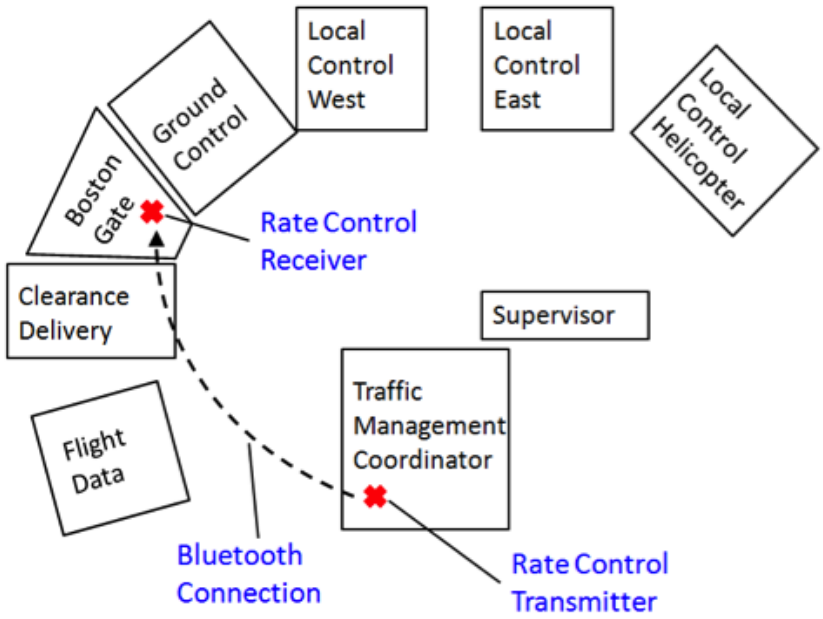
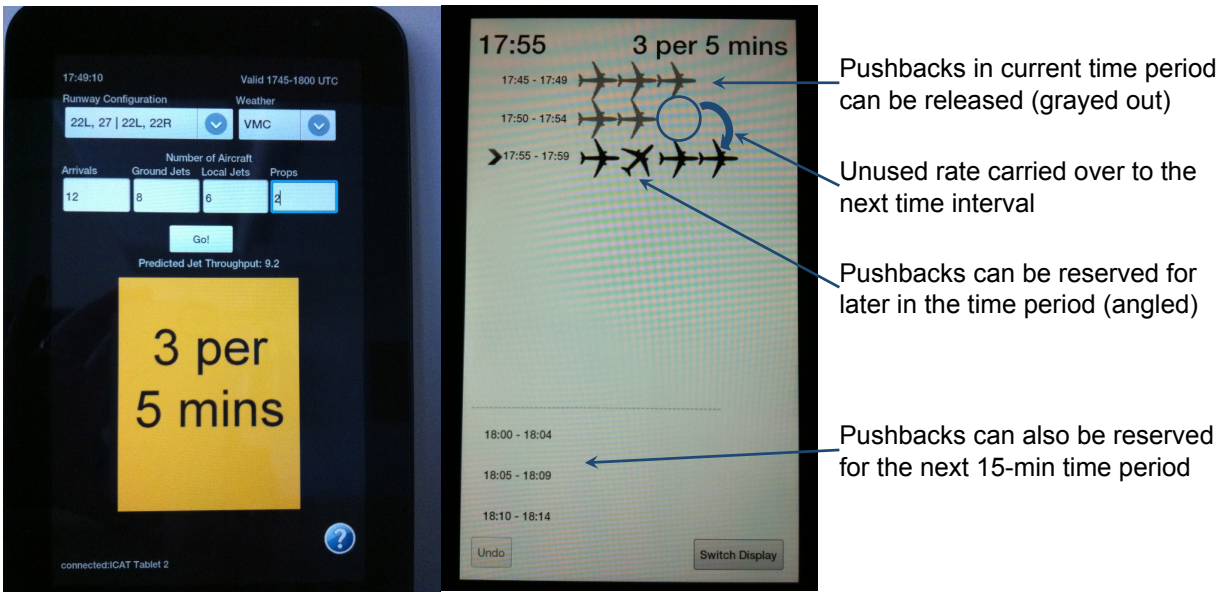


Figure 5-7: Setup of rate control transmitter and receiver in the BOS ATCT.

5.6.1 Inputs

The application developed calculates the expected departure throughput and the recommended pushback rate using look-up tables for the PRC_v2.1 algorithm. The previously defined state variables are given as inputs: runway configuration, weather (visibility conditions), expected arrival rate in the next 15 minutes, jets on ground control, jets on local control, and number of props taxiing out. The input interface is shown in Figure 5-8a.



(a) Rate control transmitter, showing the input interface.

(b) Actions in the volume control display of the rate control receiver.

Figure 5-8: The two tablets used during the 2011 field-trials at BOS.

5.6.2 Outputs

Once the suggested pushback rate is determined and transmitted, the receiver conveys the information to the BG controller through one of two display modes: the rate control and the volume control displays.

Rate control display

In this mode, the output is simply an image of a color-coded pushback rate, showing the number of allowed pushbacks per interval of minutes. With this display mode, the BG controller keeps track of the time intervals and the number of aircraft that have already pushed back. When the demand for pushbacks exceeds the recommended rate, an aircraft is held until the next time interval starts.

Again, the BG controller has to keep track of holding the aircraft and then releasing them when the next time interval begins.

Volume Control Display

This display mode helps BG controllers keep track of the number of aircraft that had called and been released. It was observed during the field trials in 2010 that many controllers used handwritten notes to keep track of the number of aircraft released, so as not to exceed the recommended rate [110]. The volume control mode helps them with this task, and also provides visual cues of the timeline and upcoming actions.

On the volume control display, the 15-minute time period is broken down into smaller time intervals, based on the rate. For example, if the rate is 3 per 5 minutes, the display shows three rows of three aircraft icons, with each row corresponding to a 5-minute time interval. The current time interval is indicated by a small black arrow to the left of the time interval. Aircraft can only be released during an ongoing time interval. Other positions can only be reserved. Any unused release spots for a given time interval roll over to the next time interval. The following actions are available in the volume control display (illustrated in Figure 5-8b):

1. **Releasing an airplane**: If a flight calls for pushback, one of the aircraft icons in the ongoing time interval is selected. The color of the icon changes from black to gray, indicating that it has been released.
2. **Reserving an airplane**: If a flight calls for pushback and there are no more positions available in the current time interval, the BG controller tells the aircraft to hold and reserves a position for it in a future time interval. This is done by selecting an aircraft icon on the display, which then rotates by 45 degrees to indicate that it has been reserved. When that aircraft is eventually released, the controller clicks on the aircraft icon again; the icon then rotates back and turns gray.
3. **Reserving a position in a future time period**: An aircraft position for an upcoming 15-minute time period can be reserved by clicking on the white space next to that time period. A rotated aircraft icon then appears in order to indicate a reservation. When the appropriate time period arrives and the suggested rate has been calculated, that aircraft icon will appear already reserved.

5.6.3 Tablet deployment

During the 2011 field trials, a member of the research team gathered and inputed data into the rate control transmitter. The rate control receiver was located next to the BG controller, who chose between rate control display and volume control display. It is expected that in an actual deployment, the traffic management coordinator (TMC) or the Supervisor would collect and input the data. In half of the test hours, the BG position was staffed by an individual controller, and in the other half, it was merged with another position – Clearance Delivery (CD) or the TMC (Figure 5-7). The merging of positions was conducted to investigate the potential implementation of PRC without requiring an additional controller at BG, which is typically only functional during times of extreme weather.

5.7 Field trials evaluation

Although the PRC strategy was tested at BOS during 19 demo periods between July 18th and September 11th 2011, there was very little need to control pushbacks when the airport operated in its most efficient configuration 4L, 4R | 4L, 4R, 9, or when demand was low. In only eight of the demo periods was there enough congestion for gateholds to be experienced. A total of 144 flights were held, with an average gate-hold duration of 5.3 min. During the most congested periods, up to 44% of flights experienced gateholds.

Table 5.2: Summary of gate-hold times for the eight demo periods with significant gateholds.

Date	Period	Configuration	No. of gateholds	Total gateholds (min)
7/18	4.45-8PM	22L, 27 | 22L, 22R	14	28
7/21	5.15-9PM	22L, 27 | 22L, 22R	42	384
7/22	5.15-8.30PM	22L, 27 | 22L, 22R	50	290
7/24	5.15-8PM	4L ,4R | 4L, 4R, 9	12	13
7/28	5.30-8PM	4L ,4R | 4L, 4R, 9	7	13
8/11	5.30-8.15PM	22L, 27 | 22L, 22R	6	9
8/14	5.00-6.30PM	22L, 27 | 22L, 22R	1	1
	6.30-7.30PM	4L ,4R | 4L, 4R, 9	0	0
9/11	5.30-6.30PM	4L ,4R | 4L, 4R, 9	0	0
	6.30-8.15PM	22L, 27 | 22L, 22R	12	23
Total			144	761

5.7.1 Congestion control

In this section, we describe the basic results of the PRC field-tests with regard to congestion control.

Two illustrative examples

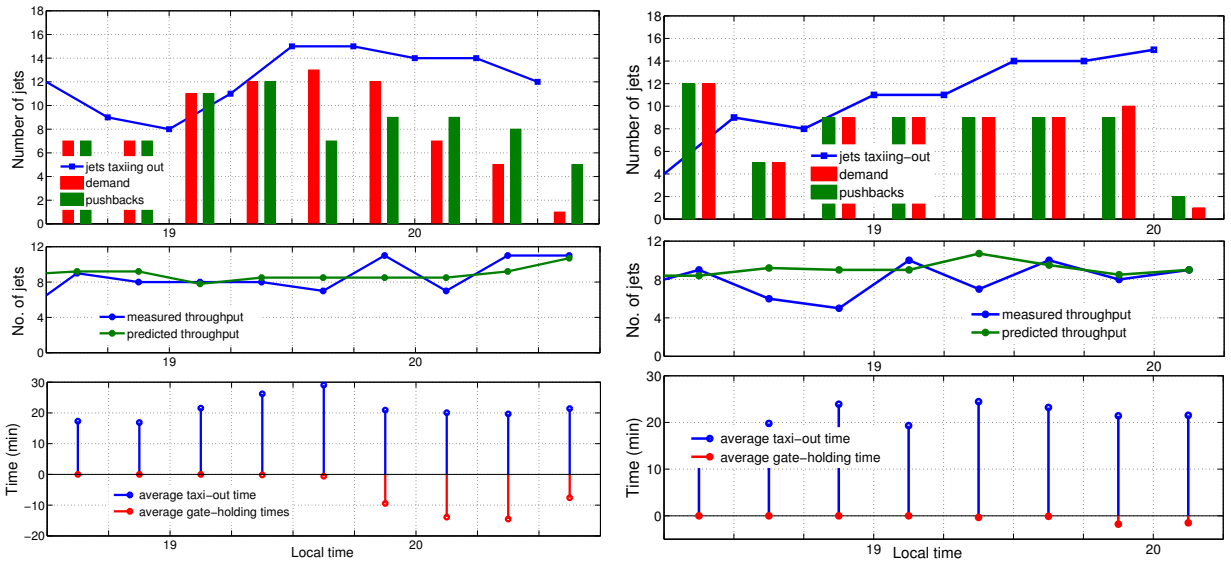
Here, we examine a day with significant gateholds (July 21, 2011) and a day with few gateholds (September 11, 2011) Figure 5-9 depicts the events of the demo period on July 21, 2011 and September 11, 2011 divided into 15-minute windows. The top plots show the demand for pushbacks (that is, the number of aircraft that called for push), the pushbacks that were cleared, and the resulting number of jet aircraft actively taxiing out. The center plots show the throughput predicted by our algorithm and the throughput measured using ASDE-X data. Finally, the bottom plots show the average taxi-out times and gate-holding times for aircraft that pushed back in each time interval.

From the top plot in Figure 5-9a, we observe that as the number of jet aircraft taxiing-out increases and exceeds 14, gateholds are initiated in order to regulate the traffic to the desired state. For configuration 22L, 27 | 22L, 22R, the desired state is 13-14 aircraft on the surface. We see that the algorithm reduces this number, from 15 to 14, and then to 12.

The airport stays in the desired state despite the high variance of the departure throughput (middle plot of Figure 5-9a) and the rounding-off of the recommended pushback rates. An objective of the PRC_v2.1 algorithm is to balance congestion management with predictability (and thus ease of implementation), and this is done fairly well. While the desired traffic level stays within 1 or 2 units of the target value, the recommended pushback rate does not fluctuate excessively, and stays centered around 8 aircraft per 15 minutes throughout the high-demand period, 1930 to 2030 hours.

With regards to the predictability of the pushback control strategy, we also note that the traffic level at the airport was successfully regulated to a similar extent during the high-demand period (1930 to 2030 hours) on all days of the field trials despite the different demand patterns, departure throughput, and the duration and number of gateholds.

Finally, we show the same plots for September 11, 2011, a day without significant metering, in Figure 5-9b. On this day, a configuration change was initiated at 1835 hours. In its aftermath, the throughput was lower than predicted for the time window 1845-1900 hours. However, the pushback demand was low and the airport stayed within the desired traffic level, that is 13 to 14 aircraft for the next hour, despite non-predicted fluctuations in the departure throughput. In particular, the drop in departure throughput performance in the time window 1915-1930 hours was due to



(a) July 21, 2011

(b) September 11, 2011

Figure 5-9: Surface congestion at each time window, demand and pushbacks during each time window (top), departure throughput measurements and predictions (center) and average taxi-out times and gateholds (bottom) during each time window for two days of field-testing

wrong estimation of the departure readiness for 3 flights with EDCTs by the TMC. We note, that there were no gateholds at the time. The pushback rate was set at at 9 AC/15 min for the time period 1900-2000 hours. However, the demand followed the same trend, thus very few gateholds were necessary. The few gateholds were mainly applied for smoothing the demand evenly in this hour. By 2000 hours, some traffic built up similarly to July 21, 2011. However, the demand was subsequently extremely low and no further gateholds were necessary. Comparing Figures 5-9a and 5-9b, we note that the traffic on the surface and the taxi-out times follow similar trends (12-15 AC on the ground, and approximately 20 minutes taxi-out times) despite the departure throughput fluctuations. However, on July 21, the high pushback demand resulted in extensive gateholds so as to keep the airport in the desired operational regime. This was not necessary on September 1, when the demand was low and steady throughout the evening departure push and subsided after 2000 hours.

Runway utilization

A key objective of the field-test was to maintain pressure on the departure runways, while limiting surface congestion. By maintaining runway utilization, it is reasonable to expect that gate-hold times translate to taxi-out time reduction. We confirm that runway utilization was not impacted

by the control strategy by validating that the runway queue was always loaded with at least one aircraft. This validation was performed both visually during the trials, and by Khadilkar using ASDE-X data, as described in [110].

5.7.2 Translating gate-hold times to taxi-out time reductions

Having field tested the pushback rate control protocol, the next step is to quantify the benefits of the approach. The main dimensions of the benefits that we address are the taxi-out time and fuel burn reductions. Intuitively, it is reasonable to use the gate-hold times as a surrogate for the taxi-out time reduction, as long as runway throughput is maintained. We test this hypothesis through a simulation of operations with and without metering.

Simulation set-up

The purpose of these simulations is to estimate the taxi-time savings and to investigate the fairness of the strategy in terms of the distribution of gateholds. In particular, we compare three different sets of outcomes:

1. *Data from actual operations:* This case corresponds to the system behavior during the pushback rate control demo periods. The taxi-times and queuing times are measured using ASDE-X data.

2. *Simulation predictions:* This case corresponds to the simulated output of pushback rate control demo periods. In this simulation, flights are cleared for pushback at the same times that they received pushback clearance (after being assigned gateholds) during the demo.

3. *Hypothetical (no pushback rate control) simulation:* Finally, the model is used to simulate what would have happened if pushback rate control was not in effect, that is, if flights had been cleared for pushback as soon as they called ready to push. In the simulations, the pushback clearance times for flights are set to be equal to the call-ready times, that is, all gate-hold times are set to zero.

The common elements in all simulations are the following:

1. The departure slots are fixed and determined by the data from actual operations for each day. This reflects the fact that there are differences in runway performance across days due to factors not related to the pushback rate control strategy.

2. The flights with EDCTs and DSPs are assumed to have fixed departure times-same as the ones observed in real operations. This is because these flights have pre-defined departure times.

The first step is to determine the unimpeded taxi-out times of flights using ASDE-X data, adopting the procedure outlined in Section 4.3.1. Given the pushback clearance time, the unimpeded taxi-out time and a taxiway congestion component, each flight is propagated to the runway, where it is assigned to the next available departure slot for that time period, which determines the predicted wheels-off time. The difference between this wheels-off time and the pushback clearance time is the expected taxi-out time.

The fixed departure slots are a reasonable assumption as long as there is a nonzero queue at the departure threshold. The total and mean taxi-out times from the actual data and the model predictions are expected to be the same, since the pushback times and departure slots are the same for both cases. The additional comparison of the actual and predicted runway queuing times would reflect the ability to predict the travel time from the ramp to the runway queue, and subsequently to compare the impact of the control strategy using the simulations.

The results are summarized in Table 5.3 for the two days with significant gateholds. The results pertain to flights that were released for pushback between 1675 and 2045 hours, that is, near and during the metering period. There were 21 flights with EDCTs and DSPs on July 21 and 17 such flights on July 22. As can be seen in Table 5.3, the mean taxi-out time and the mean queuing time (the time an aircraft spends in the departure runway queue) are generally predicted very well by the model.

Table 5.3: Effect of gate-holding on mean taxi-times and queue lengths.

		Actual operations			**Model predictions**		**No pushback rate ctrl.**	
Date	No. of flights	Taxi-out time (min)	Queuing time (min)	Gate-hold time (min)	Taxi-out time (min)	Queuing time (min)	Taxi-out time (min)	Queuing time (min)
7/21	121	16.5	5.7	368	16.5	5.8	19.5	7.9
7/22	121	17.9	7.2	279	17.9	7.4	20.2	9.2

In the top part of Figure 5-10 we show the instant actual and simulated queue on July 21. They match very well, so the actual queue is predicted accurately by the simulations. In the bottom part of the same figure we compare the simulated queues of July 21 with and without PRC. The evident difference between the simulated queue sizes shows the benefit of the pushback rate control strategy.

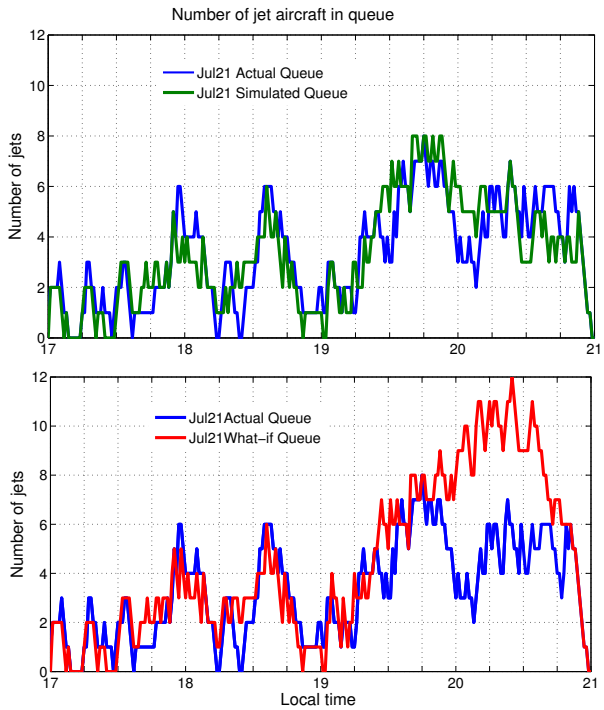


Figure 5-10: Queue sizes measured and predicted per minute, on July 21, 2011.

In addition, we conduct a benefits analysis of the fuel burn savings by using the simulated taxi-out time savings times as a first-order estimate of the actual taxi-out time savings using the methodology outlined in prior work [70, 110]. The total fuel savings are estimated to be 2,650 US gallons, which translates to average fuel savings per gate-held flight of about 57 kg.

Distribution of benefits

Equity is an important factor in evaluating potential congestion management or metering strategies. The PRC approach, as implemented in these field tests, invokes a First-Come-First-Serve (FCFS) policy in clearing flights for pushback. One would therefore expect that there would be no bias toward any airline with regard to gateholds incurred, and that the number of gateholds for a particular airline would be commensurate with the contribution of that airline to the departure traffic during the congested periods. However the taxi-out time saving predicted by the simulations is not equal to the gate holding time of each individual flight. Thus, the taxi time savings of each carrier can differ from the total time flights of this carrier were held at the gate as can be seen in Figure 5-11. This is because the benefit of holding a flight at the gate can spill over to other flights as well as explained in earlier studies of N-Control [107]. In short there are two main reasons for

this:

- **Overtaking:** Consider a scenario in which aircraft A calls for push and is authorized to push. Aircraft B calls for push just a few seconds later and is held at the gate for 3 minutes. Subsequently, aircraft B pushes back and finds itself in the departure queue behind aircraft A. In absence of the metering program, aircraft B would have pushed a few seconds and not 3 minutes later than aircraft A. This might have been enough time for aircraft B to overtake aircraft A if it taxied at a faster speed, its gate was located closer to the runway, or it performed the pushback process faster. Thus, in the counterfactual scenario, aircraft B could have ended up in front of aircraft A in the departure queue. In such a scenario the cost and the benefit of the program could swap between two flights.

- **Re-scheduling despite same sequencing:** Consider a scenario in which all pushback slots have been utilized and there are 5 more minutes left until the end of the current time-window. In this time-window 4 new flights call for pushback, each one a minute apart from the previous one. When the next time window commences, the rate is set to 3 every 5 minutes. The controller authorizes the first 3 aircraft to push back together, and the forth one five minutes later.

As a result of these two phenomena and their combinations, the benefits of the metering scheme can be divided and allocated between flights in an unpredictable manner. However, the first come first serve sequence is maintained and in general the gateholding times would be approximately equal to the taxi-time reduction experienced by each airline, as can be confirmed from Figure 5-11. However, the actual fuel burn benefit also depends on its fleet mix. Figure 5-11 shows that while the taxi-out time reductions are similar to the gateholds, some airlines (for example, the ones denoted Airlines 4, 13, 21 and 27) benefit from a greater proportion of fuel savings. These airlines are typically ones with several Heavy aircraft during the evening times.

5.7.3 Departure throughput prediction

As explained in Section 5.5, we use the algorithm PRC v_2.1 for predicting the jet departure throughput. Because of the sources of inaccuracy in both ASDE-X and ASPM data [110], we validate the predictions during shadow testing (June 30-July 17 2011) by means of visual observations and subsequently use them during the 19 days of the trials to predict the throughput. Table

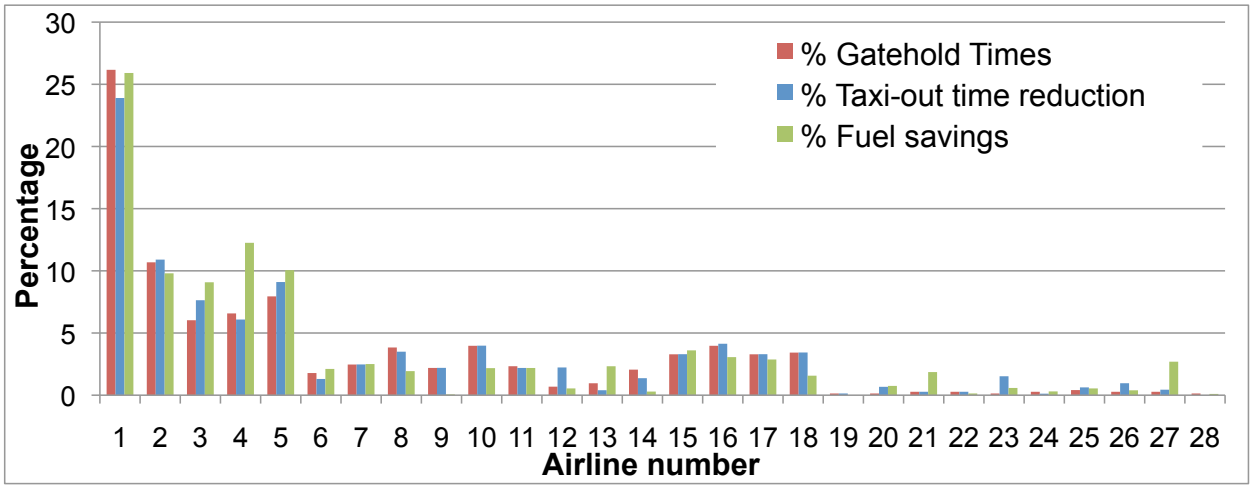


Figure 5-11: Percentage of gate-hold times, taxi-out time reduction and fuel burn savings corresponding to each airline.

5.4 reports the average error, average absolute error and root mean square error of the predicted throughput (relative to observed throughput) during 182 15-min periods of field testing.

For completeness, the corresponding errors of alternative prediction methods which we could have used are also shown:

- Predictions from PRC v_2.0, that is, the queuing model with the "unconditional" service time distribution of each runway configuration in the evenings. This algorithm would input the number of aircraft traveling and queuing into the queuing model to predict the throughput without using arrivals and props demand information.

- Predictions from the demand curves (DC), that is, using Figure 5-1 for each runway configuration to predict the departure throughput based on the total number of departing jets taxiing out.

- Predictions from regression trees (RT): This method would use trees such as the one of Figure 5-6 to predict the departure throughput given the number of aircraft traveling, queuing, props and arrival demand information.

Finally, we also compare the errors for the 93 periods where the traffic was 10 aircraft or more, because these are the times when gateholds are most likely.

Table 5.4 shows that the regression tree based prediction algorithm used in PRC_v2.1 predicts the takeoff-rate reasonably well: The mean absolute error is only 1.14 during medium and high traffic conditions (10 jets or more). However, there is little benefit from using the conditional

Table 5.4: Comparison of the estimator used and three alternatives for predicting jet departure throughput.

	All traffic conditions			$\geq$ 10 jets taxiing		
Estimator	ME	MAE	RMSE	ME	MAE	RMSE
PRC v_2.0	-0.09	1.24	1.62	0.08	1.14	1.54
PRC v_2.1	-0.20	1.25	1.64	-0.03	1.14	1.58
DC	0.71	1.32	1.74	0.64	1.18	1.69
RT	0.64	1.35	1.78	0.59	1.19	1.78

throughput distributions. By using the unconditional evenings throughput distribution, we could achieve the same, or even better prediction accuracy. While this could imply that the parametrized distributions are an artifact of over-fitting, Figure 5-6 captures an underlying trade-off between jet departure rates, props departure rates and arrival rates. We therefore hypothesize that the small size of the training dataset, or the few test days lead to high prediction errors. Another possible reason for the large variance is that we do not account for some significant hidden variables, such as summer convective weather. The model was trained using mostly non-convective days (November 2010- June 2011), but it was applied during the months of July and August which are subject to high convective activity. In particular, at 53 out of the 182 time windows experienced significant convective weather in the North-East US.

More importantly, we note that the prediction algorithm accuracy is in agreement with the uncertainty considered in the design of the pushback control strategy. For configuration 22L, 27 | 22L, 22R and when at least 10 jet departures were taxiing, the highest underestimation of the departure throughput was 2.7. The algorithm tries to maintain a queue of at least 3 aircraft for this configuration, as explained in Section 5.5.4. Similarly, for configuration 4L, 4R | 4L, 4R, 9 and when at least 10 jet departures were taxiing, the highest underestimation of the departure throughput was 3.7. For this configuration, the algorithm tries to maintain a queue of at least 5 aircraft. The above observations suggest that the inventory targeted by the algorithm at the queue was set at the correct level in terms of avoiding runway underutilization; a more aggressive congestion control policy could have resulted in an empty runway queue in these two cases. However, a reduction in the variance of the actual or predicted departure throughput could lead to more aggressive control of the traffic. The importance of a sufficient inventory at the runway queue has also been noted by other researchers [103].

The demand curve based model (DC) and the regression trees (RT) has worse jet departure

throughput predictions than the other two models, and tend to overestimate the throughput. These models are trained with ASDE-X data, which underestimates the traffic levels because of the delay between the actual pushback and ASDE-X capture times [110]. A purely statistical predictive model therefore yields high errors reflecting ASDE-X measurement errors.

5.7.4 Evaluation of the Decision Support Tool

A survey of the controllers was conducted to gather their opinions on the study as a whole, and specifically on the implementation and use of the tablet. The survey was presented to the controllers after the field-tests had been completed. There are 21 respondents in total, 15 of whom were BG in 2010, 13 in 2011, and 12 during both years.

We solicited quantitative ratings on five topics: Whether they thought fuel burn decreased, whether surface traffic flows improved, whether throughput was adversely impacted, whether the new (tablet) display was easier to use that the color-coded cards used in 2010, and whether they found the new display easy to use. The histograms of the results are shown in Figure 5-12. We see that the survey responses were generally positive, and that the controllers liked the new tablet displays as well. We also hypothesize that there may have been some confusion about the scale on the question of throughput, since several of the controllers who agreed that the throughput was adversely impacted also agreed that the surface traffic flow improved.

Thirteen responses were also positive about combining BG and another position. Ten of these responses suggested CD, three indicate the TMC, and one each indicate GC and Flight Data (more than one position could be indicated). The survey also shows that the controllers like the tablet volume control display format a lot. Among the comments on the best features are: “the ability to touch planes”, “reserve spots”, “count the planes and account for aircraft with long delays”, “allows me to push & tells me to hold”, and “easy to use & understand”. Suggestions for improvement include increasing the icon sizes and maintaining more pressure on the runway. Finally, the controllers are satisfied with the modifications between 2010 [110] and 2011 field trials, with one of them remarking: “Liked the improvement in just one year”.

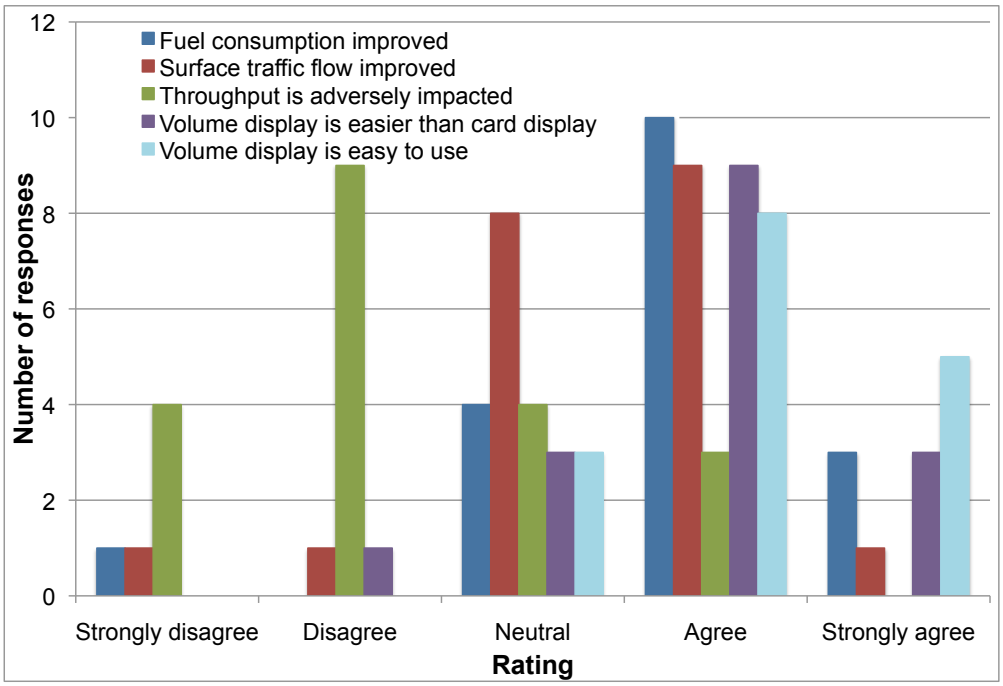


Figure 5-12: Histogram of responses from air traffic controller survey regarding PRC at BOS.

5.7.5 Qualitative observations

Compatibility with traffic flow management initiatives

An important goal of this study was to investigate the compatibility of PRC with traffic flow management initiatives. Under highly convective weather, the abundance of these programs leads to many target departure times, schedule disruptions or flight cancellations. As a result, congestion does not build up, and there is no metering.

However, there are days during which the traffic management programs do not lower demand significantly. July 18, 2011 was one such day. There were two Minutes-In-Trail (MINIT) programs during the departure push of this day: All westbound flights had 5 MINIT between 2245 and 2335 hours, and 3 MINIT between 2335-0030. At the same time, there was a 5 MINIT restriction for all flights over LUCOS. These programs spread out the departures, and decreased the opportunities for metering, but did not lower the overall departure demand. This resulted in a combination of the MINIT programs and the congestion metering program between 2245 and 2300 hours. The integration of the two programs was very simple and effective: The total number of flights released per time window was set by the metering program, and the mix by the MINIT program. For example, if the pushback rate were 3/5 min while westbound flights had 5 MINIT, the controller

would release two flights with no MINIT restrictions along with a westbound departure. Similarly when the pushback rate was 4/5 min, the controller would release three flights with no MINIT restrictions along with a westbound departure.

The field tests also showed that the approach is capable of handling target departure times (e.g., EDCTs), but for that it is preferable to get EDCTs while still at gate. Flights with EDCTs were generally exempt from gateholds. However, on days in which the BG and TMC positions were merged (for example, July 21), delays due to the controlled departure times could be absorbed as gateholds. During the July 21 demo period, two flights with EDCTs called for push when gateholds were in effect. The controller informed them that gateholds are in effect, asked them to hold their push and called the appropriate centers to obtain their controlled departure times. Subsequently, he released them from their gate so that they could takeoff at their assigned times. Both flights took off a minute before their assigned times. In this way, the flights with EDCTs absorbed their delays at the gate, saved fuel, and were integrated with the rest of the traffic after pushback clearance. This made it easier for the controller to handle them and ensure that they met their controlled departure times.

Increased predictability

An additional benefit of the approach is the ability to communicate expected pushback times to pilots in advance. For instance, on July 21, more than 10 aircraft were on hold at the beginning of the periods 2000-2015 hours and 2015-2030 hours. Once the suggested pushback rate was given to the controller at the start of each time period, the controller communicated the expected release times to all aircraft on hold. These flights received their release times several minutes in advance, which could be useful in planning ground resources.

Natural metering effect

The suggested pushback rate in very low congestion time-periods is 1 per min. However, we noticed that the merging of the BG position with another position resulted in a natural rate of 1/min without explicit gateholds. For example, when the BG position was merged with the TMC, after the controller cleared an aircraft that called for push, he/she would have to spend the rest of the minute for a traffic management task (such as, calling the center to obtain an EDCT). As a result, the next aircraft would only be released after a minute, resulting in a natural metering of 1 per min unless a lower rate was recommended.

This effect offers a good opportunity for the operational deployment of a metering scheme at no added personnel cost. The gate position could easily be merged with another position, such as Clearance Delivery or the TMC.

5.8 An alternate model for the runway service process

5.8.1 Motivation

In Section 5.3.4, we proposed a model for the runway service process assuming that the service times follow an Erlang distribution. This enabled us to model the evolution of the queuing system using the set of Equations (5.5) - (5.16). However, this model presents several challenges. First, it is hard to parametrize. For example, if we want to modify the shape of the Erlang distribution so as to observe the impact of the coefficient of variation of the service times on the control strategy, we need to numerically solve the Equations (5.5) - (5.16) for the new service time distribution $(k', k'\mu)$ and subsequently run the policy iteration algorithm to obtain the optimal policies. In other words, we cannot utilize a previously derived solution at any stage of the process. In addition, for larger values of the shape parameter, the dimensionality of the system becomes prohibitive. Moreover, the state of the embedded Markov chain maps to the state of the system only in an approximate manner as different *stages of work* have to be mapped to a single state (Equations (5.31) - (5.32)). Finally, if we have multiple throughput distributions, conditioned on external variables, the corresponding embedded Markov chains will not necessarily be in the same state space. For example, if we have two throughput distributions, with shapes k_1 and k_2, where $k_1 \neq k_2$, the space of the chain of the queuing system corresponding to the first distribution will be $\{0, 1, \ldots, k_1 \cdot C\}$ and the second $\{0, 1, \ldots, k_2 \cdot C\}$. Transitioning from the throughput distribution of the former to that of the latter can be done only by approximately mapping the probabilities of states $\{0, 1, \ldots, k_1 \cdot C\}$ to those of $\{0, 1, \ldots, k_2 \cdot C\}$. For these reasons, we chose to develop the approximate algorithm PRC_v2.1 to make use of the conditional throughput forecasts instead of introducing the throughput forecast as a state variable.

5.8.2 The $(M(t)|R_0)/D_s/1$ model

Here, we propose an alternate model for the runway service process. We assume that during each time window the service rate is deterministic, but not known. It is sampled from a finite set $\mu_1, \mu_2, \ldots, \mu_s$ with weights derived from the empirical distribution. The set $\mu_1, \mu_2, \ldots, \mu_s$, of

cardinality s is the support of the empirical distribution function. With reference to the empirical throughput distribution of Figure 5-3, this would imply that in each 15-minute window, the throughput process is deterministic with rate 5, 6, ..., or 13. The probability of each rate equals the probability mass of each rate of the empirical throughput distribution. For example, the probability that the service rate is 10/15 min is 0.277. This model yields an exact match of the empirical and modeled probability distributions of the departure throughput. On the downside, it does not model the fact that a service rate of 10 AC/15 min does not imply uniformly spaced service times of 1.5 min.

We summarize that notation used in this section:

- μ_i: Deterministic service rate ($AC/15\,\text{min}$).
- M: Set, of cardinality s, of all deterministic service rates $\mu_1, \mu_2, \ldots, \mu_s$.
- $(R, Q; \mu_i)$: State of the queuing system given the deterministic service rate μ_i.
- F: Set, of cardinality z, of all throughput distribution forecasts $f_1, f_2, \ldots, f_z$.
- $w(i; f_j)$: Probability that the service rate equals μ_i given throughput forecast f_j.

Assume for now a single deterministic service rate, μ, which implies $M = \{\mu\}$. In a given time window, the system resembles a transient $M(t)/D/1$ queuing system with the exception that the number of arrivals during a time interval is known. Following the framework of Section 5.3.4, we denote it as $(M(t)|R_0)/D/1$. For analyzing this system we extend the framework proposed by Koopman [73]. In this framework, the service epochs are a priori marked on the time axis. Continuing the previous example, where the service rate is assumed to be 10/15, the (potential) service time epochs are marked at times $1.5, 3, \ldots, 15$ minutes from the beginning of the time window. This assumption implies that if an aircraft arrives at an empty system in minute 1, it will wait until minute 1.5 before its service starts. Thus, delays at lower states might be overestimated. On the other hand, this assumption makes the analysis of the system tractable: If at epoch 0, R_0 aircraft are taxiing out, the probability mass function $\hat{g}$ of k arrivals between the departure runway service times i and $i+1$ assuming that $j-k$ aircraft have already arrived, is:

$$\begin{aligned}
\hat{g}(i,j,k) &= \Pr\{k \text{ arrivals in}(t_i, t_{i+1}]|(R_0-(j-k))\text{arrivals in}(t_i, \Delta]\} \\
&= \begin{cases} \binom{R_0-(j-k)}{k}\left(\frac{\tau_{i+1}-\tau_i}{\Delta-\tau_i}\right)^k\left(\frac{\Delta-\tau_{i+1}}{\Delta-\tau_i}\right)^{(R_0-j)}, & \text{if } 0 \le k \le j, j \le R_0, \tau_{i+1} \le \Delta \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{5.33}$$

Analogously to the $(M(t)|R_0)/E(k)/1$ system, the state of the runway system is denoted as $(R, Q; \mu)$, where R is the number of aircraft traveling (aircraft to be delivered to the queuing system), Q is the number of aircraft in the queuing system (in service, or in queue) and μ is the assumed deterministic service rate. Now, we observe that $\hat{g}(i, j, k)$ is the probability of transitioning from state $(R_0 - (j-k), j-k+\mathbf{1}_{\{j-k\geq 1\}}; \mu)_{\tau_i}$ to state $(R_0 - j, j; \mu)_{\tau_{i+1}}$. The condition $j - k \geq 1$ implies that there were one or more aircraft in the system before the arrival of the k aircraft between service times i and $i+1$, and one of them was served. At epoch 0 the system is in state (R_0, Q_0, μ). The state of the queuing system at time Δ, $\hat{Q}_\Delta(\mu)$, is a probabilistic function of the initial value (R_0, Q_0, μ), the functions $\hat{g}(i, j, k)$ describing the probability of each allowable transition, and the assumed service rate μ.

In the actual system, for each throughput forecast f, we have a finite set $\mu_1, \mu_2, \ldots, \mu_s$ of deterministic service rates each with probability $w(1; f), w(2; f), \ldots, w(s; f)$. The state of the queuing system at Δ, Q_Δ is given by the weighted sum of the $\hat{Q}_\Delta(\mu_i)$'s:

$$Q_\Delta(f) = \sum_{i=1}^{s} w(i; f) \cdot \hat{Q}_\Delta(\mu_i) \tag{5.34}$$

Equation (5.34) reveals the benefit of this formulation: The probability vector of the state of the system $Q_\Delta(f)$ given a throughput forecast f is decomposed in a weighted sum of $\hat{Q}_\Delta(\mu_i)$'s, which are independent of the weights of the summation $w(1; f), w(2; f), \ldots, w(s; f)$. Thus, a different throughput distribution f can be modeled by simply changing the weights $w(i; f)$ in Equation (5.34). We denote this queuing model of deterministic service times sampled from a finite set and a known number of arrivals at random times as $(M(t)|R_0)/D_s/1$. Moreover, Equation (5.34) offers the ability to track each individual arrival at the queue. Each possible transition is assigned a probability ($\hat{g}(i, j, k)$) and a cost. The cost has two components, the queuing and the non-utilization of the runway. For the queuing cost we have that the first out of k arrivals between service times i and $i+1$ will encounter a system with $j-k$ aircraft, the second $j-k+1$, the kth a system with $j-1$. Thus, each transition can be explicitly penalized in terms of its expected queuing delay (in minutes). Similarly, each transition from an empty system ($j-k=0$), can be penalized in terms of loss of runway utilization. A loss of runway utilization can be also expressed in minutes: It is the minutes of additional delay later flights are likely to incur because of the capacity loss.

5.8.3 Comparison of the two models

The first step is to compare the performance of the two models presented, the $(M(t)|R_0)/E(k)/1$ and the $(M(t)|R_0)/D_s(t)/1$ in terms of predicting the state of the queue after a 15-minute period given the range of possible initial conditions G and D, for the example considered in Section 5.5. In this example, we have one throughput forecast: The jet departure throughput at BOS configuration 22L, 27 | 22L, 22R in evening times.

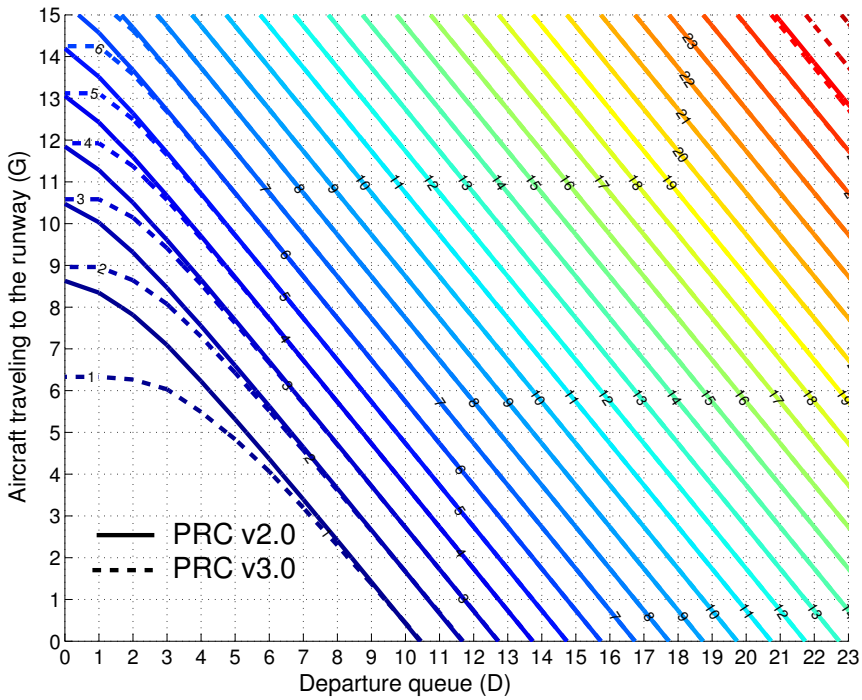


Figure 5-13: Expected queue length after 15 minutes as a function of the number of aircraft in the departure queue (D) and the number of aircraft traveling to the runway (G) for the two models, $(M(t)|R_0)/E(k)/1$ in solid line, $(M(t)|R_0)/D_s(t)/1$ in dashed line.

We note that the two models predict for most initial conditions the same value for the expected queue, except for the states, where the initial value of the queue is very low (0-3 aircraft). For example, given 0 aircraft queuing and 12 aircraft traveling to the runway, $(M(t)|R_0)/D_s/1$ predicts an expected queue of 4 after 15 minutes, whereas $(M(t)|R_0)/E(k)/1$ predicts an expected queue of 3. We conjecture that $(M(t)|R_0)/D_s/1$ underestimates the throughput (and thus overestimates the queue) compared to the $(M(t)|R_0)/E(k)/1$ in these occasions, because of the simplifying assumption that the service times are a priori equally spaced in the time window.

For deriving optimal pushback policies we obtain the optimal cost-per-stage solutions for the

$(M(t)|R_0)/D_s/1$ model. We have for the probability of the queue being in state q after Δ minutes:

$$p_q(r,q) = \sum_{i=1}^{s} w(i;f) \cdot \Pr(\hat{Q}_\Delta(\mu_i) = q) \tag{5.35}$$

The Bellman equation for the average optimal cost-per-stage, c^*, for the control of the $(M(t)|R_0)/D_s/1$ model can be written as :

$$c^* + h^*(r,q) = \min_{\lambda \in \Lambda} \big\{ (\bar{c}(r,q) + \mathbf{p}_q(r,q) \cdot \mathbf{h}^*(\lambda) \big\} \tag{5.36}$$

for $r \in \{0, 1, \ldots, \lambda_{max}\}$ and $q \in \{0, 1, \ldots, C\}$

We note that the state space is significantly reduced from that of the $(M(t)|R_0)/E(k)/1$ (Equation (5.30)). The state space describing the $(M(t)|R_0)/D_s/1$ system is $(1 + \lambda_{max}) \cdot (1 + C)$, and it does not change with the parameter s, whereas the state space describing the $(M(t)|R_0)/E(k)/1$ system is $(1 + \lambda_{max}) \cdot (1 + k \cdot C)$ and grows linearly with k.

We name the resulting strategy PRC_v3.0. The cost of capacity loss experienced with an empty queuing system is set equal to 120 minutes of queuing delay. This is because at an airport like BOS, a minute of loss of runway utilization can delay subsequent departures cumulatively for as long as 60 minutes during the evening departure push. We multiply this delay by 2, because the delay resulting from loss of runway capacity directly increase takeoff delays, in contrast to queuing delays which do not change takeoff delays in general, but where they are allocated (gate vs. surface).

Analogously to Figure 5-5, we show in Figure 5-14 the scatterplot between the optimal pushback rate $\hat{\lambda}_\tau(G_\tau, D_\tau)$ and the expected $\bar{D}_{\tau+\Delta}(G_\tau, D_\tau)$, for all $0 \leq G \leq \lambda_{max}$ and $0 \leq D < C$, along with a fitted convex non-increasing function that minimizes absolute deviations from the data for the strategy PRC_v3.0. We also show the scatterplot and the fitted function for PRC_v2.0. We observe that the two strategies deviate when the expected queue length is equal to zero. In this case, PRC_v3.0 recommends one fewer aircraft than PRC_v2.0. At higher states, the small deviation is due to the approximate derivation of the optimal policies for PRC_v2.0 from the embedded chain (Equations (5.31) and (5.32)). Another reason for the differences between the two policies is the different cost structure considered: PRC_v2.0 penalizes queuing states, whereas PRC_v3.0 penalizes expected queuing delays.

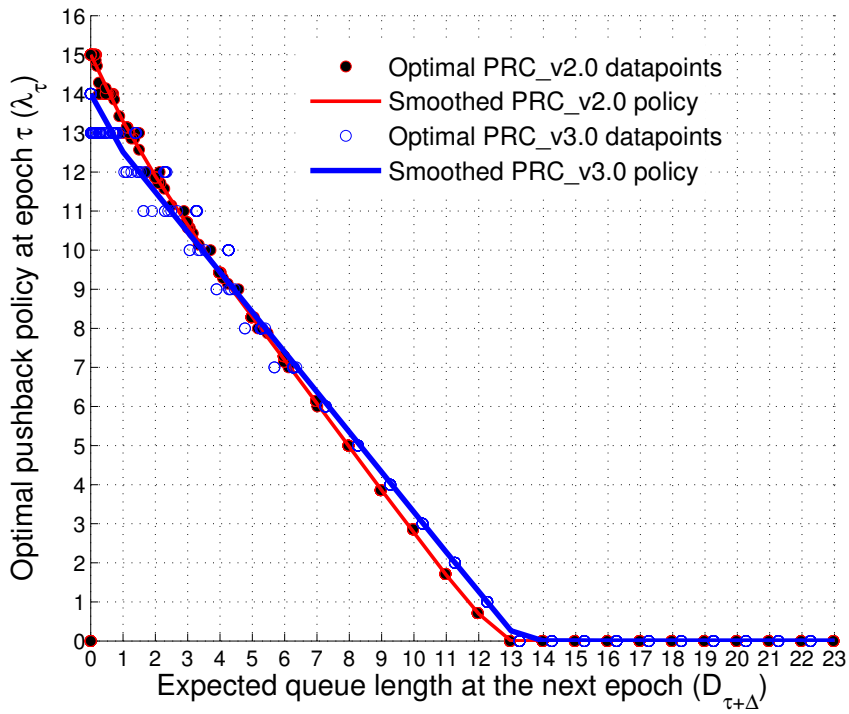


Figure 5-14: Optimal pushback policy as a function of the expected queue $\bar{D}_{\tau+\Delta}$ at the next epoch $(\tau + \Delta)$ for the policies PRC_v2.0 and PRC_v3.0.

5.8.4 Parametric analysis

As a next step, we modify the algorithm PRC_v3.0 in order to investigate the validity of the approximate PRC_v2.1. For this, we modify the state space to include the throughput forecast (F) as a state of the system:

$$N_t = (G_t, D_t, F_t) \tag{5.37}$$

In this case, the average optimal cost per stage, c^* is calculated as:

$$c^* + h^*(r, q, \mathrm{f}) = \min_{\lambda \in \Lambda} \big\{ (\bar{c}(r, q, \mathrm{f}) + \sum_f p_f \mathbf{p}_q(r, q, \mathrm{f}) \cdot \mathbf{h}^*(\lambda, f) \big\} \tag{5.38}$$

where p_f is the probability of each throughput forecast and $\mathbf{p}_q(r, q, \mathrm{f})$ is the probability vector of the state of the queue at the end of the time-window given that the state at the beginning of the time window is (t, q, f). We call the resulting algorithm PRC_v3.1.

The formulation is to determine the average optimal cost per stage, c^* for PRC_v3.1 in its general form is:

$$c^* + h^*(r,q,\mathrm{f}) = \min_{\lambda\in\Lambda}\Big\{(1-\sum\beta_i-\sum\gamma_i)[\bar{c}(r,q,\mathrm{f})+\sum_f p_f\mathbf{p_q}(r,q,\mathrm{f})\cdot\mathbf{h}^*(\lambda,f)]$$
$$+\sum\beta_i[\bar{c}(r+i,q,\mathrm{f})+\sum_f p_f\mathbf{p_q}(r+i,q,\mathrm{f})\cdot\mathbf{h}^*(\lambda-i,f)]$$
$$+\sum\gamma_i[\bar{c}(r-i,q,\mathrm{f})+\sum_f p_f\mathbf{p_q}(r-i,q,\mathrm{f})\cdot\mathbf{h}^*(\lambda+i,f)]\Big\} \tag{5.39}$$

In Figure 5-15, we show the optimal pushback policies for each conditional forecast shown in Figure 5-6.

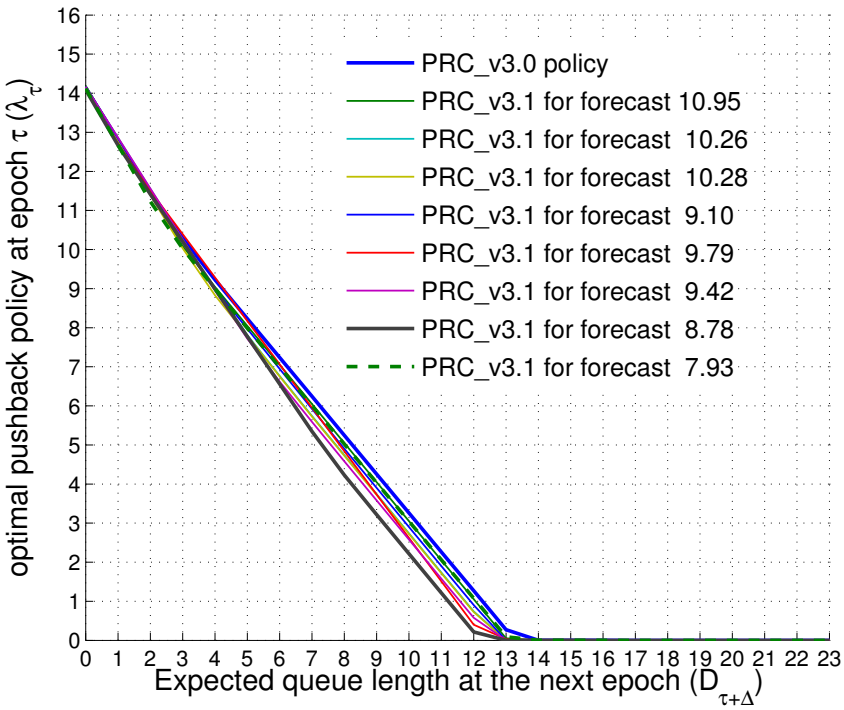
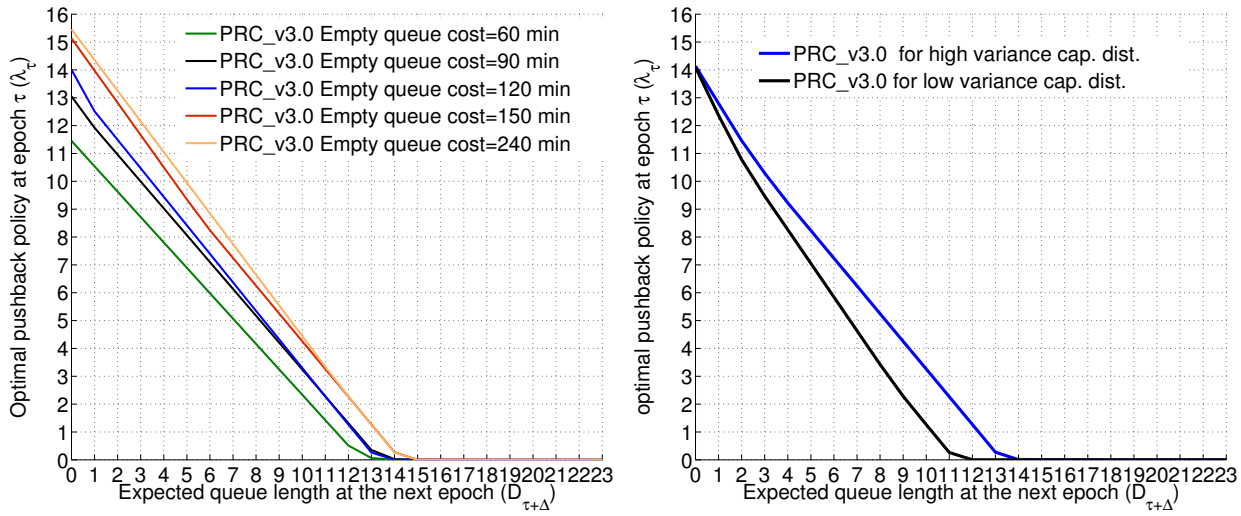


Figure 5-15: Optimal pushback policy $\hat{\lambda}_\tau$ as a function of the expected queue $\bar{D}_{\tau+\Delta}$ at the next epoch $(\tau+\Delta)$ for the policies PRC_v3.0 and PRC_v3.1.

We notice that the policies of PRC_v3.0 are more "conservative" compared to those derived using the conditional throughput forecasts. This was expected because PRC_v3.0 is derived using a throughput distribution of higher variance. The biggest difference between PRC_v3.0 and PRC_v3.1 policies is for throughput forecast 8.8. This forecast has the lowest variance (Figure 5-6) and the algorithm can afford to be more aggressive when the queue is expected to be longer than 6 AC.

We note, that the expected value of the throughput forecast does not impact how different its curve from PRC_v3.0 lies. The impact of the expected value of the throughput is accounted by

the expected queue length. This was indeed the rational for using the expected queue length as a quasi-state of the system when proposing PRC_v2.1. We observe this also in Figure 5-15. The forecast with the lowest throughput (7.93) is shown with the dashed line and its optimal pushback policy is practically the same as that of PRC_v3.0. This analysis shows that PRC_v2.1 is an effective approximate algorithm that allows us to incorporate on-line any conditional, or updated throughput forecast (as long as the general distribution of the departure throughput does not change significantly). The updated throughput forecast could be for a special occasion pertaining to a particular 15-minute interval which is not part of the training data. For example, the TMC might decide that due to route closures, the throughput during a 15-minute interval will be only 5 AC/15 min. PRC_v2.1 can be used for calculating the optimal pushback rate for this occasion despite the fact that such a throughput forecast is not a member of the set of throughput forecasts, as shown in Figure 5-6.



(a) Optimal pushback policy $\hat{\lambda}_\tau$ as a function of the expected queue $\bar{D}_{\tau+\Delta}$ at the next epoch $(\tau+\Delta)$ for different cost functions.

(b) Optimal pushback policy $\hat{\lambda}_\tau$ as a function of the expected queue $\bar{D}_{\tau+\Delta}$ at the next epoch $(\tau+\Delta)$ for a high variance throughput distribution and a low variance throughput distribution.

Figure 5-16: Sensitivity analysis examples.

We can also use this model to perform sensitivity analyses. An example of such sensitivity analysis is shown in Figure 5-16a, where we investigate the impact of the cost function on the optimal policy derived with PRC_v3.0 (and shown in Figure 5-14). For this analysis, the cost of an empty queuing system is set equal to a range of values between 60 and 240 minutes of queuing delay. As expected, a higher cost associated with an empty queue results in more conservative

pushback rate policies, which aim at securing a larger inventory of aircraft at the departure queue. For example, when the cost of an empty queuing system is 60 min the policy targets a queue of 2 AC, when it is 90 min, at 3 AC, and for a cost of an empty queuing system at 240 min, it targets a queue of 6 AC. As already explained in Section 5.5.4, for our chosen empty queuing system cost (120 min), the algorithm targets a queue of 3-4 AC.

Finally, we use this model to more closely investigate the impact of the variance of the departure throughput on the optimal policies. We hypothesize that, similar to other operations management problems [114], the reduced variance of the departure throughput would result in less *inventory* in the runway queue, and thus more *aggressive policies*. For this investigation, we compare the policy PRC_v3.0 developed for the departure throughput of Figure 5-3 (*high variance*) to the policy for one with the same average throughput (9.74), but where all the probability mass is concentrated at 9 and 10. In other words, in this *low variance* scenario, the service rate in a 15-minute time period is 9 AC/15 min with probability 0.26 and 10 AC/15 min with probability 0.74. In Figure 5-16b, we show the PRC_v3.0 optimal pushback policies for the *low variance* and the *high variance* throughput distributions. As expected, PRC_v3.0 is more aggressive for a throughput distribution with lower variance.

5.8.5 Simulation of PRC at PHL

PHL departure model

Following the process outlined in Chapter 4, we model the departure process for the major runway configuration of PHL, 26, 27R, 35 | 27L, 35, which was in use 74% of the time under VMC in 2011. In this runway configuration, the major departure runway is 27L, and the major arrival runway, 27R. According to ASDE-X data analysis presented in Section D.3 of the appendix, there is one departure on runway 35 for every 11.5 departures on 27L. Additionally, the thresholds of the two runways are very close to each other, and the aircraft heading to both of them are part of the same flow. The detailed results of the model are discussed in Appendix G. We also present the operational throughput envelopes for this configuration in Appendix C.

We can use the model to predict the evolution of the departure throughput and taxi-out times over a day at PHL. In the upper plot of Figure 5-17, we show the average number of pushbacks and the average number of takeoffs (or departures) that was recorded during each 15 minute time window for all days in which this runway configuration was in use in 2011. We also show the average

number of departures of this runway configuration of PHL, as predicted by the model. In the lower part of Figure 5-17, we show the actual and predicted average taxi-out times for the flights that pushed back in each 15-minute time window. We observe that the model is representative of an average day at PHL.

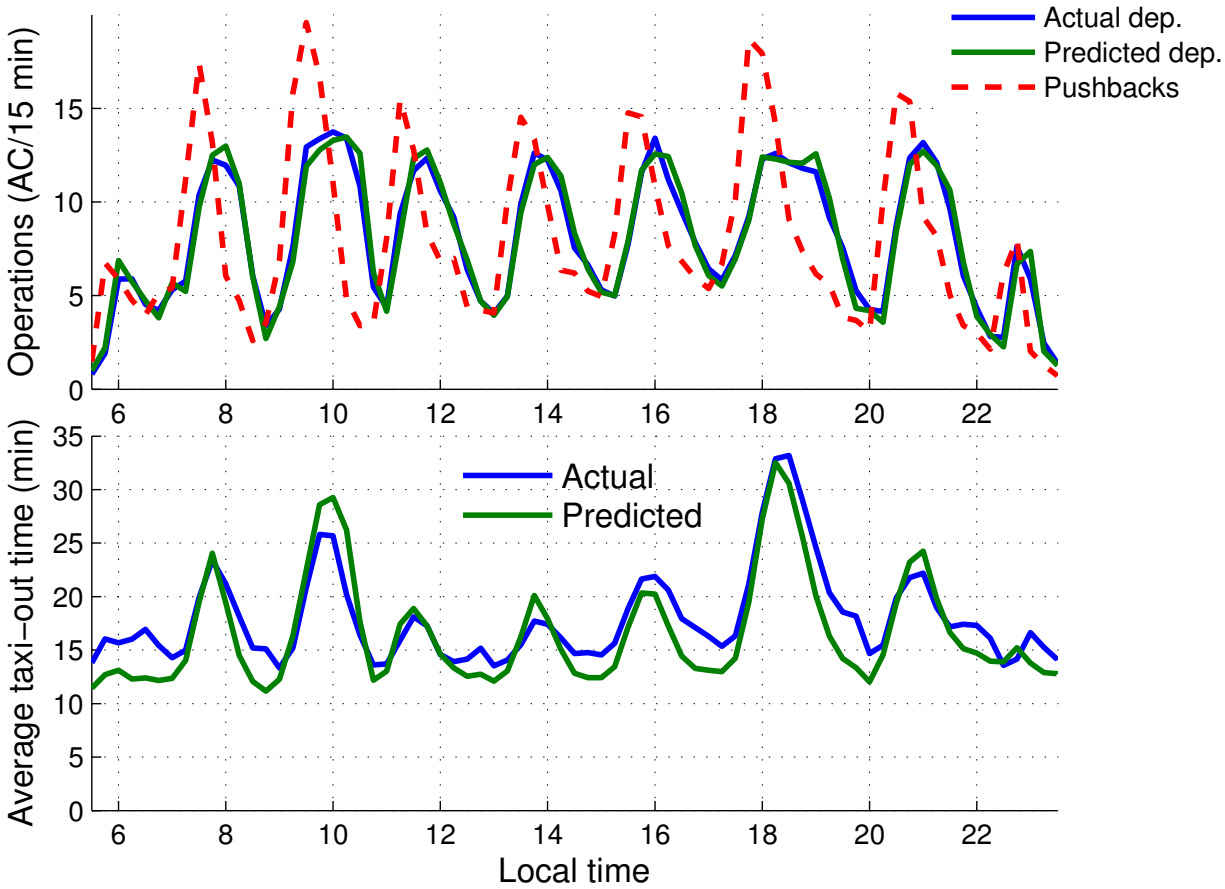


Figure 5-17: Average number of pushbacks, and average numbers of actual and predicted takeoffs by time of day at PHL in 2011 (top); Average actual and predicted taxi-out times (bottom).

PRC_v3.1 calibration

The modeling effort described in Appendix G provides the empirical capacity distributions, thus the μ and w parameters of Equation (5.35) that are necessary for estimating the queuing transition probabilities are known. Additionally, as already explained in Section 5.8.4, it is not necessary to recalculate the transition probabilities, nor the congestion costs. They are simply weighted with the appropriate weights w's in the dynamic program (Equation (5.39)). The maximum pushback rate is set to 24 AC/15 min, which is the maximum pushback rate observed currently at PHL

after eliminating outliers. This rate is assumed to be the a natural threshold for the maximum admissible rate of arrivals into the departure process (pushbacks).

The next step involves the decision of the optimal time window, Δ, for this runway configuration. For this, we do a simple flow analysis. On average, in a controlled scenario, aircraft enter the system at the same rate as they exit, that is the average departure throughput (13.0 AC/15 min). The average unimpeded taxi-out time is 11.8 min and parameter α is 0.22 min/AC. From this, we calculate that the average number of aircraft traveling from the gates to the departure runway during each minute is 12.6 AC. Thus, the average travel time of each aircraft entering the system is $11.8\,\text{min} + 0.22 \times 12.6\,\text{min} = 14.6\,\text{min}$. Therefore, the time window Δ is set at 15 min.

The last step involves calculating the probabilities $\beta_i(R, \lambda)$ and $\gamma_i(R, \lambda)$, which are necessary for deriving the system dynamics (Equation (5.27)). For this calculation, we use Monte-Carlo simulations to estimate the empirical distribution of the number of aircraft traveling to the runway at the next epoch, $R_{\tau+\Delta}$ given the current number of aircraft traveling, R_τ, and the current pushback rate, λ_τ. We let the system evolve starting from a randomized initial condition, and we select random pushback rates every 15 minutes. These rates are allocated to airlines according to their relative presence at the airport. Every 15 minutes, we record the transition $R_{\tau+\Delta}$, given the current R_τ and pushback rate λ_τ. Finally, we derive β_i and γ_i from the simulated empirical distributions $R_{\tau+\Delta} = g(R_\tau, \lambda_\tau)$.

Simulation setup

By simulating PRC_v3.1 strategy at PHL, we are interested in observing the following:

1. PRC represents the state of the surface of the airport with only two variables: the number of aircraft traveling to the departure queue (G) and the number of aircraft in queue (D). For computational tractability, we do not have further information on where the traveling aircraft are distributed. A simulation will illustrate the price of this state space reduction.

2. PRC_v3.1 assumes the approximate model $(M(t)|R_0)/D_s/1$ for the runway service process. The simulation will indicate if this model is adequate.

3. PRC can be compared to other popular control mechanisms. For this comparison, we use *N-Control*, the most popular state-dependent control mechanism, and *Slot-Control*, the most popular non-feedback control applied in airport environments.

In Appendix G.1, it is seen that the airport saturates when 20 aircraft are taxiing-out. Thus, $N^* = 20$. For calibrating N-Control[1], we use an aggressive choice of N_{ctrl}, that is $N_{ctrl} = N^* = 20$. Simulations show that for $N_{ctrl} = 20$, aircraft do not incur additional delay resulting from gate-holding. Given that $N_{ctrl} \geq N^*$, the resulting taxi-out time reduction is the highest that can be achieved with N-Control.

For simulating Slot-Control, we use the results of Section G.2. There, we show that the average departure capacity of this runway configuration at PHL is 13AC/15 min. It is also explained that the average departure capacity does not change significantly with arrival throughput. A straightforward way to simulate Slot-Control, is therefore to limit pushbacks to the departure capacity, that is 13 AC/15 min. This limitation is imposed as a cap on the total number of pushbacks allowed in a 15-minute interval, and not as a pushback rate. Slot-Control is the simplest control policy; it is open-loop and straightforward to implement. Its performance offers a benchmark for the additional value provided by feedback-control policies, such as N-Control and PRC.

For all control policies, we impose the additional constraint that the pushback rate cannot exceed 4 AC/min, which was the maximum number of pushbacks/minute achieved at PHL in 2011-after eliminating outliers. This additional constraint aims at incorporating the practical limitation that pushback coordination and communication require a certain minimum time. This phenomenon relates closely to the natural metering effect observed at BOS, as discussed in Section 5.7.5.

Finally, the earliest possible pushback time of each flight is its recorded actual pushback time. This means that in all scenarios, pushbacks can only be delayed, and not advanced. In addition, for the cases of PRC and Slot-Control, if the pushback requests in a 15-minute time window are fewer than the optimal pushback rate and the pushback cap respectively, the remaining slots are unutilized. The likely resulting loss of runway utilization is because of the absence of sufficient demand at this time period, and not because of the control scheme. Similarly, for the case of N-Control, there are instances without sufficient pushback requests to bring the number of aircraft on the surface to the N_{ctrl} value.

Simulation results

For simulating the three strategies, we run 100 Monte Carlo simulations sampling the unimpeded taxi-out time of each flight from the corresponding distribution, and using a displaced exponential service time at the runway (Monte Carlo simulation setting 4, explained in Section 4.5.2). We also

[1] An overview of N-Control is provided in Section 1.2.4.

simulate a *do-nothing* scenario. The results are summarized in Table 5.5, for a total number of 136,430 flights that pushed back and departed in this configuration at PHL in 2011. The column "mean delay" lists the additional takeoff delay that flights incur as a result of the control scheme. It is calculated by subtracting the take-off time in the *do-nothing* scenario from the takeoff time in the controlled scenario.

Table 5.5: Taxi time predictions for PHL from simulating different control strategies.

Control algorithm	Mean taxi-out time (min)	Taxi out time st. dev. (min)	Mean delay (min)	Mean holding time (min)	Number of flights held
Do-nothing	18.46	8.53	0.00	0.00	0
N-Control	16.85	5.82	0.00	1.61	31,325
PRC	16.83	5.86	0.03	1.66	28,594
Slot-Control	17.03	6.60	0.27	1.70	52,042

From Table 5.5, we note that the PRC simulation shows that the strategy works as designed: It manages to reduce average taxi-out times by 1.66 minutes, while adding a very short average delay of 0.03 minutes. We believe that this a result of the $(M(t)|R_0)/D_s/1$ model used for its derivation. As explained in Section 5.8.3, this model underestimates the risk of an empty departure queue at low values of demand. PRC also reduces the variation of taxi-out times. The N-Control strategy results in slightly smaller taxi-out time savings, but at zero added delays. It is noteworthy that the PRC strategy achieves very similar results to the N-Control strategy despite being applied periodically, that is, every 15 minutes. We conjecture that this is because of the predictive nature of PRC. Instead of aiming to keep the taxiing-out traffic below 21 AC, we use information on the current state of the airport and predict the departure capacity and the departure queue in the next 15 minutes. The pushback rate is then set so as to optimize the load of the queue. These findings suggest that the optimal PRC policies result in achieving taxi-out time savings very close to those of N-Control.

From Table 5.5, we also notice that a Slot-Control policy performs significantly worse. Its taxi-out time savings are less than those of both N-Control and PRC, and are achieved by delaying flights by 0.27 min on average. This would imply a total added delay of 614 hours over the course of a year. The reason behind this weaker performance is that the departure process in PHL is very stochastic. The shapes of the fitted Erlang distributions are 1-3, as shown in Section G.2. Clearly, an open-loop control policy is not well suited for such a dynamic process. Despite the fact that the number of pushbacks is capped at the average departure capacity of the system, loss of runway utilization occurs often enough that this capacity loss propagates to delay later aircraft.

We also remark that Slot-Control would result in more variable taxi-out times than N-Control and PRC. Taxi-out times grow much higher under Slot-Control because of the absence of a feedback mechanism at times of significant congestion.

In Figure 5-18, we show a visualization of the averaged traffic statistics by time of day resulting from a single run of N-Control simulation at PHL in 2011. Similarly, in Figure 5-19, we show a visualization of the averaged traffic statistics by time of day resulting from a single run of PRC simulation. From the lower plots of the figures, we notice that both strategies are very effective in reducing long taxi-out times, and in particular removing the taxi-out time peaks at 1000 and 1900 hours. By comparing the upper plots of Figures 5-18 and 5-19, we notice that the controlled pushback rate exhibits a similar trend under both strategies. It is high in the beginning of each departure push, but it is subsequently rapidly reduced. This is because both strategies aim at initially loading the runway queue, and subsequently regulating the flow of aircraft on the surface.

We also note that because of the differences between the two strategies, the pushback rates at the beginning of each departure push are always slightly higher under the PRC strategy (for example at 1745 hours). As discussed in Section 5.5, at low traffic conditions, the optimal optimal pushback rate under PRC is very high, since it aims to build up the queue at the runway. Subsequently, given the current state of the surface, (G, D), and the predicted capacity, the pushback rate is regulated so as to maintain a desired inventory of aircraft at the queue, in this case 5-6 aircraft. This means, that on average, and in steady state, there will be 5-6 aircraft in queue and 13 aircraft taxiing to the runway. Although the initial level of traffic is higher for PRC than for N-Control, it subsequently stabilizes at lower values (18-19 AC) on average.

In Figure 5-20, we show the averaged traffic statistics by time of day resulting from a single run of the Slot-Control simulation at PHL in 2011. We notice that the trend of pushbacks under Slot-Control is very different from that under the two other strategies. The pushback rate is always capped at 13 AC/15 min. For example, the evening departure push is evenly distributed in the 1-hour time window 1745 - 1845 hours. From the lower plot of Figure 5-20, we observe that the smoothing of the pushbacks results in significant taxi-out time reduction. Aircraft pushback at the same rate as they takeoff, and delays build up very slowly. This behavior is in agreement with the literature and operational experience [30, 93].

In Figure 5-21, we contrast the simulated taxi-out times from the three control strategies to the do-nothing approach during the evening times. We notice that in the primary evening departure push, between 1730 and 2000 hours, all control strategies achieve significant taxi-out time reduc-

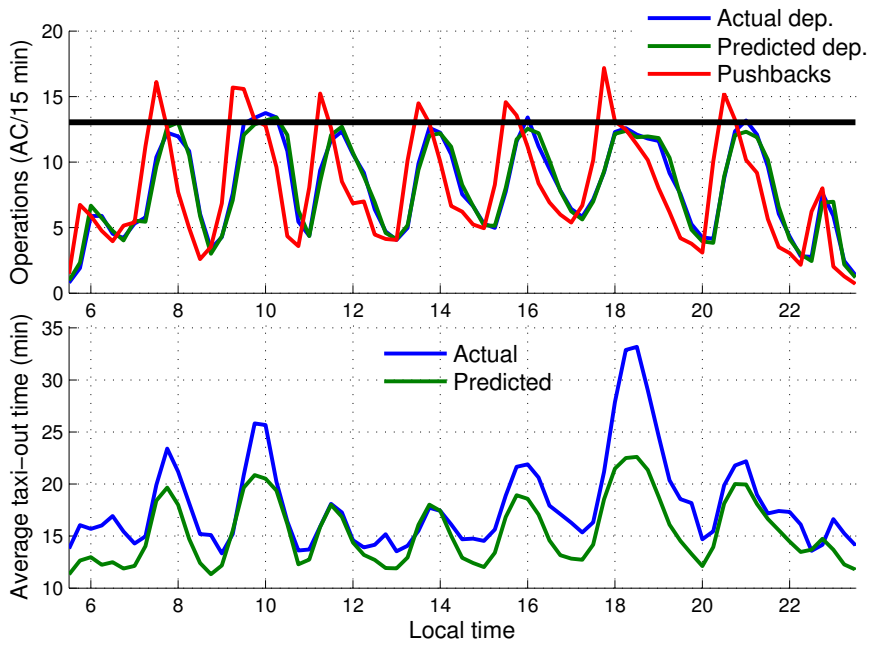


Figure 5-18: N-Control simulation: Average departure capacity (in black), average number of pushbacks, average number of actual and simulated takeoffs at PHL in 2011 (top); Average actual and simulated taxi-out times (bottom).

tions. Under Slot-Control, taxi-out times are low at the beginning of the departure push, because aircraft push back at the same rate as the service rate (the departure capacity). However, between 1800 and 1830 hours, a significant number of Heavy aircraft push back, and the departure capacity is reduced. As time progresses and aircraft arrive at the queue at a rate greater than the service rate, queuing delays build up. By contrast, under both PRC and N-Control, delays are higher than those under Slot-Control before 1815 hours, but subsequently, and until the end of the departure push at 2000 hours, they become significantly lower than those under Slot-Control. We also notice that between 1830 and 1930 hours, PRC performs better than N-Control. This is because we predict that departure capacity will be lower than its average value due to the large number of Heavy aircraft taking off in this time frame (Figure G-7), and, with PRC, we can adjust the pushback rate accordingly. By contrast, with the N-Control protocol, we cannot use of this information.

In conclusion, we note that despite the slight increase in average delays, Slot-Control is an easily implementable strategy that, in addition to managing congestion, increases the transparency and predictability of the system. Aircraft pushback times are determined in advance. At the other end of the spectrum, N-Control holds and releases aircraft according to the congestion state of the airport. Thus, an aircraft asking for permission to push, or being on hold, has no information on

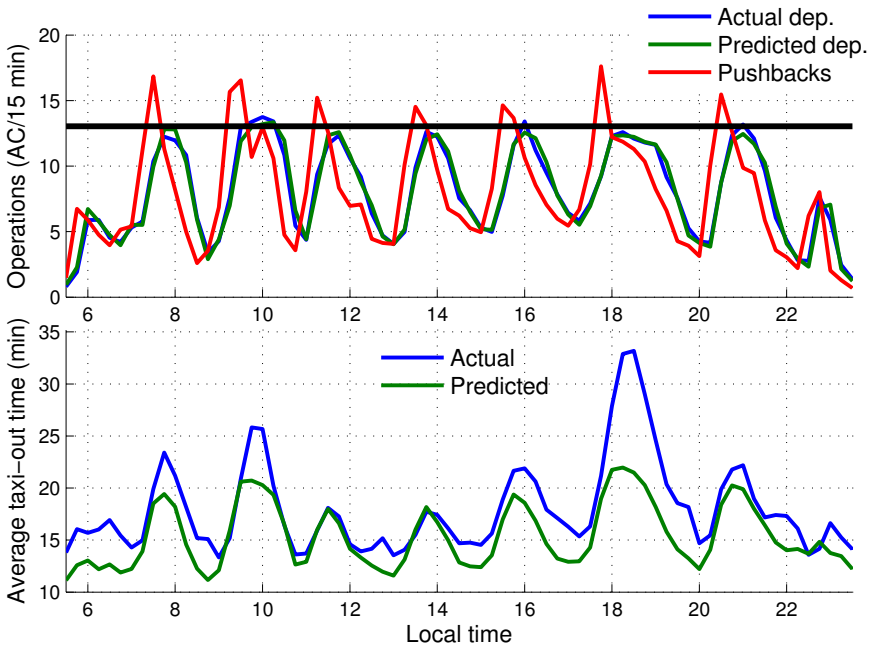


Figure 5-19: Slot-Control simulation: Average departure capacity (in black), average number of pushbacks, average number of actual and simulated takeoffs at PHL in 2011 (top); Average actual and simulated taxi-out times (bottom).

its release time. We believe that PRC offers a good compromise in the middle of this spectrum. Congestion is efficiently managed, high runway utilization is achieved, and pushback rates, or equivalently pushback tactical slots, are allocated every 15 minutes. Additionally, the pushback rate can be adjusted to accommodate changes in the departure capacity. Finally, as we confirmed from the field-testing of PRC at BOS (Section 5.7.5), PRC has the additional benefit that it can be easily combined with other air traffic flow management programs.

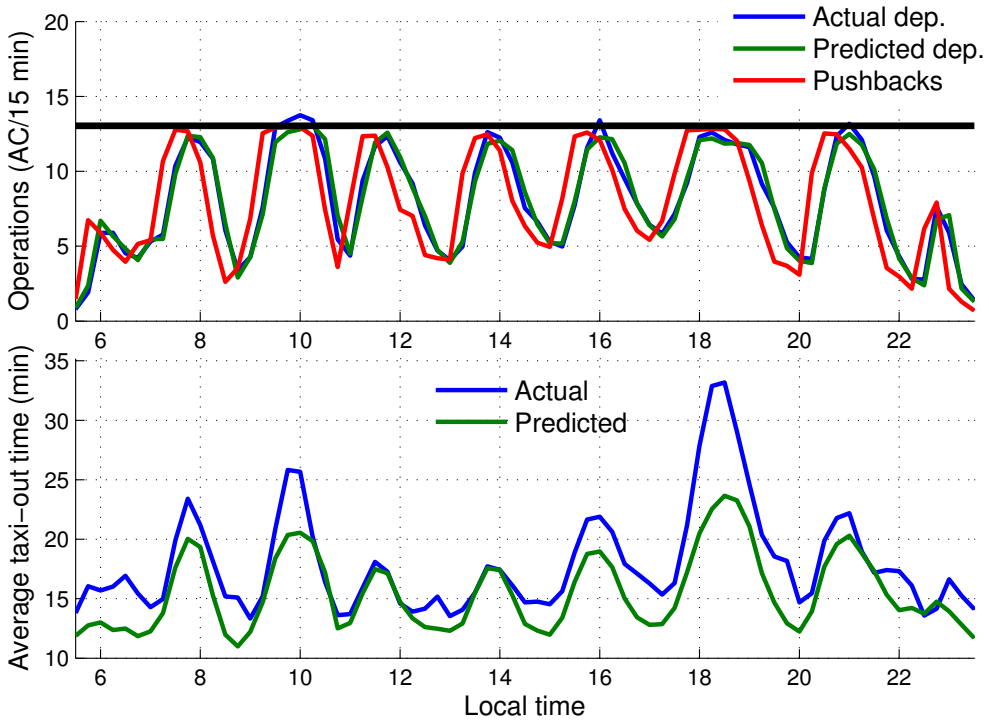


Figure 5-20: PRC simulation: Average departure capacity (in black), average number of pushbacks, average number of actual and simulated takeoffs at PHL in 2011 (top); Average actual and simulated taxi-out times (bottom).

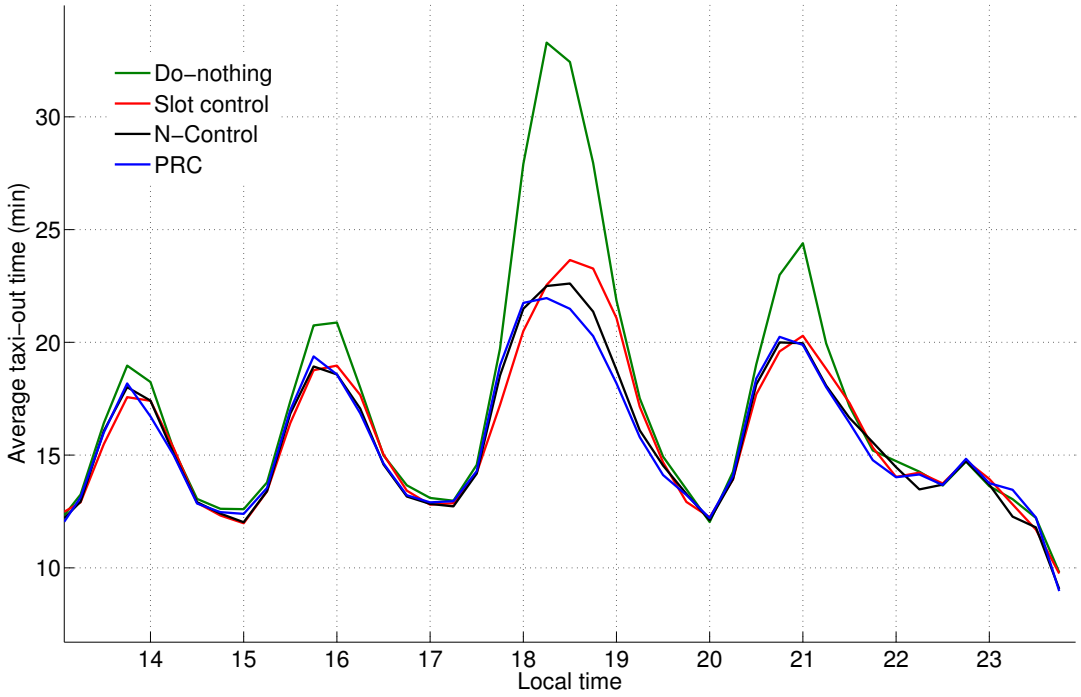


Figure 5-21: Comparison of the performance of the control strategies in the evening times

5.9 Conclusions

This chapter presented the results of the demonstration of PRC at BOS in 2011. We developed and simulated PRC algorithms using dynamic programming to balance the objectives of maintaining runway utilization and limiting surface congestion. We also developed and field-tested a decision support interface to display the suggested pushback rate, which helped the controllers keep track of requests for pushback, gateholds, and other metering constraints. During 8 four-hour tests conducted during the summer of 2011, fuel use was reduced by an estimated 9 US tons (2,650 US gallons), while carbon dioxide emissions were reduced by an estimated 29 US tons. Aircraft gate pushback times were increased by an average of 5.3 minutes for the 144 flights that were held at the gate. A survey of the air traffic controllers involved in the 2011 demo indicated support for the PRC approach, the manner of implementation, and the displays and communication protocols developed for the deployment of such strategies. Finally, a simulation of PRC at PHL airport showed that the developed approach is a promising solution for reducing congestion in a practical and predictable way.

Chapter 6

Summary and Next Steps

6.1 Summary of results

In the first part of this thesis, we developed a new method, the *operational throughput envelope*, for characterizing airport capacity and applied it to several busy US airports. We also extended the *operational throughput envelope* method to estimate the dependence of the departure throughput on parameters other than the arrival throughput and derived the following results. For the case of BOS, we showed that the departure throughout is more heavily dependent on the ratio of props in the fleet mix than the arrival throughput. At all major runway configurations of BOS, the departure throughput was estimated to decrease with an increase in arrival throughput by at most 2.6 AC/15 min, for an increase of arrival throughput from 0 to 14 AC/15 min. Increasing the number of props in the fleet mix from 0 to 5 increased the departure throughput by at least 4 AC/15 min. By contrast, at EWR, the departure throughput was found to be relatively inelastic to the fleet mix of the departing aircraft. AT JFK, the departure capacity was found to be underutilized in current operations.

We then extended the developed methodology to study interactions among the three major airports of the NY Metroplex, JFK, EWR and LGA. We found that operations at the three airports are not adversely impacted by operations at the other airports. We attributed this finding to the current airspace design, which keeps operations at the three airports separated. We finally derived capacity envelopes for the system comprising the three airports under different configurations. We estimated that the total balanced capacity of the Metroplex is 59 AC/15 min, the departure priority capacity is 53 AC/15 min, and the arrival priority capacity is 63 AC/15 min.

Subsequently, we presented an analytical model for the departure process. The model used a

stochastic and dynamic queuing model that provided estimates for the performance of the airport, mean taxi-out times, queuing delays and their variances. We used the model for predicting taxi-out times at EWR. We estimated its parameters using ASPM data from 2011 and showed that it can accurately estimate delays in 2007, the busiest year in the recent past. We also showed that the variance estimates can be used to assess the delay uncertainty of a pushback schedule.

In the last part of the thesis, we developed a Pushback Rate Control algorithm using dynamic programming to balance the objectives of maintaining runway utilization and limiting surface congestion. We field tested the proposed algorithm at BOS during the summer of 2011. We also developed a decision support interface to display the suggested pushback rate, and help the controllers keep track of requests for pushback, gate-holds, and other metering constraints. The field-tests were comprehensively analyzed, and the suggested approach showed promising results. During 8 four-hour tests conducted during the summer of 2011, fuel use was reduced by an estimated 9 US tons (2,650 US gallons), while carbon dioxide emissions were reduced by an estimated 29 US tons. Aircraft gate pushback times were increased by an average of 5.3 minutes for the 144 flights that were held at the gate. In addition, positive controller's feedback supported the feasibility of the proposed scheme. Lastly, we simulated the proposed strategy at PHL and showed that it is an effective compromise between state-dependent control and static congestion control.

6.2 Future research directions

In this section we suggest directions for future work regarding the three main components of the thesis.

Characterization and estimation of airport capacity

In the analysis presented in Chapter 2, arrival throughput is assumed to be an independent variable. To obtain a further enriched understanding, we could extend the analysis to include the arrival throughput as a function of arrival demand. Arrival demand is much harder to measure as arriving aircraft are distributed over the larger terminal airspace area controlled by the TRACON. In addition, the demand measurements can be censored. Aircraft may be on hold or delayed outside the terminal airspace, as holding patterns are usually outside the terminal airspace. TRACON also has the authority to change the aircraft acceptance rate at its airspace depending on the conditions (runway configuration in use, visibility, fixes closures). If the TRACON sets a very low acceptance

rate, we hypothesize that the arrival demand, as measured by the number of aircraft in the area surrounding the airport, is also very low. However, the underlying demand may be much higher and inbound aircraft can be delayed in other parts of the airspace, even on the surface of the airport of origin. Thus measuring arrival throughput conditioned on persistent arrival demand is a challenging and pertinent research direction. In addition, measuring inter-arrival spacing conditioned on an arrival queue *with pressure* would yield inter-arrival spacing distributions and serve as a comparison to the theoretical separation requirements. Moving forward, the dependence of arrival throughput on factors such as the arrival fleet mix, and the prevailing winds could be studied.

Model of the departure process

The departure process model proposed in Chapter 4 is a static model in the sense that it is not updated as it runs: Given a pushback schedule, it predicts operations in every time window using information only on external conditions (arrival demand, weather etc.) and not the state of the system. For tactical use of this model, that is for predicting the departure throughput and the taxi-out times in the next time-window, it would be useful to modify the model to dynamic. In this case, the user, or a decision support system could provide information such as the departure queue length, the departure throughput, or aircraft facing downstream constraints. With this, the model would update its state estimates. As it has been shown in literature, efficient dynamic updating of the state probabilities can be very challenging, yet provide significant gains [15, 104].

In addition, in the analytical models developed in this work, the departure throughput distribution was estimated using exogenous variables, that is, variables independent of the process being modeled. These variables are usually the arrival throughput and route availability. A promising topic of future research would involve modeling the departure capacity distribution using endogenous information such as the sequence of aircraft in the departure queue, or the number of Heavy aircraft that can depart in the next time-window. A dynamic and stochastic model using endogenous variables for modeling its dynamics, while still yielding closed forms solutions, is another challenging research direction. A first step towards such types of models is demonstrated in Appendix G, where information about the estimated type of aircraft taxiing out is used for predicting the departure throughput.

Dynamic control of the departure process

In the models developed in Chapter 5, we model the state of the surface of the airport with two variables, the number of aircraft traveling to the departure queue and the number of aircraft in the departure queue. A natural extension would be to include a third state for the ramp. This added level of information on the location of aircraft on the surface is hypothesized to yield more effective control strategies. At the same time, modeling the additional state in a computationally tractable manner could be challenging. Similarly, there are many airports in the US with multiple ramp towers (LGA, PHL, SEA) controlling different areas of the ramp. Another research direction would include the modification of the PRC strategy for deriving separate rates for different ramp towers. This could be done by introducing a state variable for each ramp, or by allocating the central rate to the different ramp towers in an efficient and fair manner.

One feature of the PRC strategies is that they give an optimal pushback rate for the next time window. There is a clear opportunity in using this information for enhancing the departure throughput in addition to managing congestion. For example, if the departure runway of the airport considered offers dispersal headings, one could choose the specific aircraft that push back and their sequence in order to maximize the probability of dispersal headings, or the expected throughput. Given the chosen sequence, the updated throughput forecast would be fed back into the PRC algorithm to suggest a new pushback rate which would be at least as high as the original. Naturally, this algorithm would require information on the earliest times that aircraft are available for pushback in the following time window and the ability to violate the FCFS principle.

A good candidate airport for testing such a strategy is PHL because the main departure runway 27L can be used for dispersal headings. In addition re-sequencing of pushbacks could be applied initially to the major user of the airport, US Airways. Similarly, the US Airways ramp tower houses also the US Airways Operations Center, thus reliable information on the earliest pushback times of the US Airways flights can be easily obtained.

Appendix A

Airport Diagrams

A.1 John F. Kennedy International Airport(JFK)

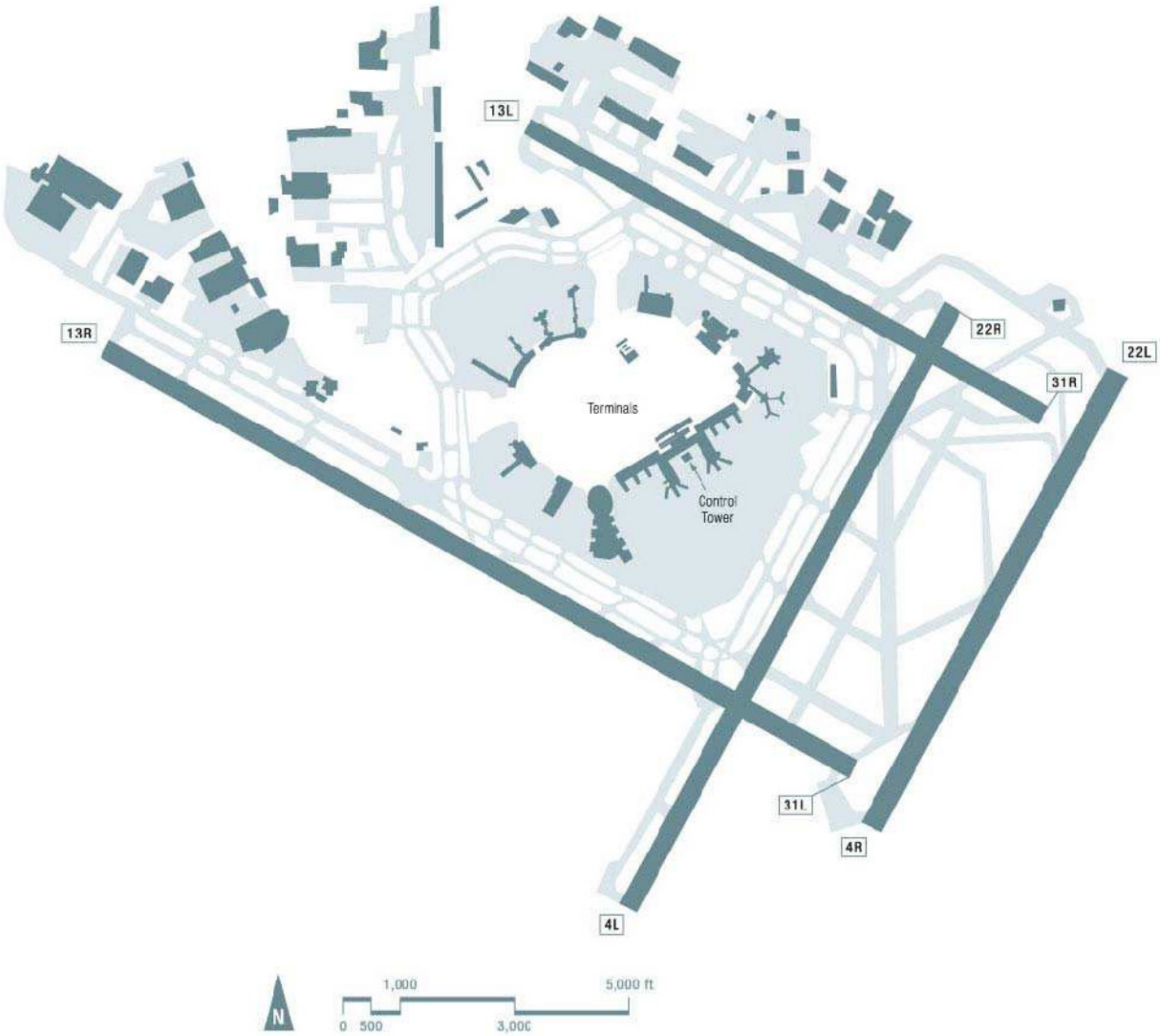


Figure A-1: JFK airport diagram[83]

A.2 Newark Liberty International Airport (EWR)

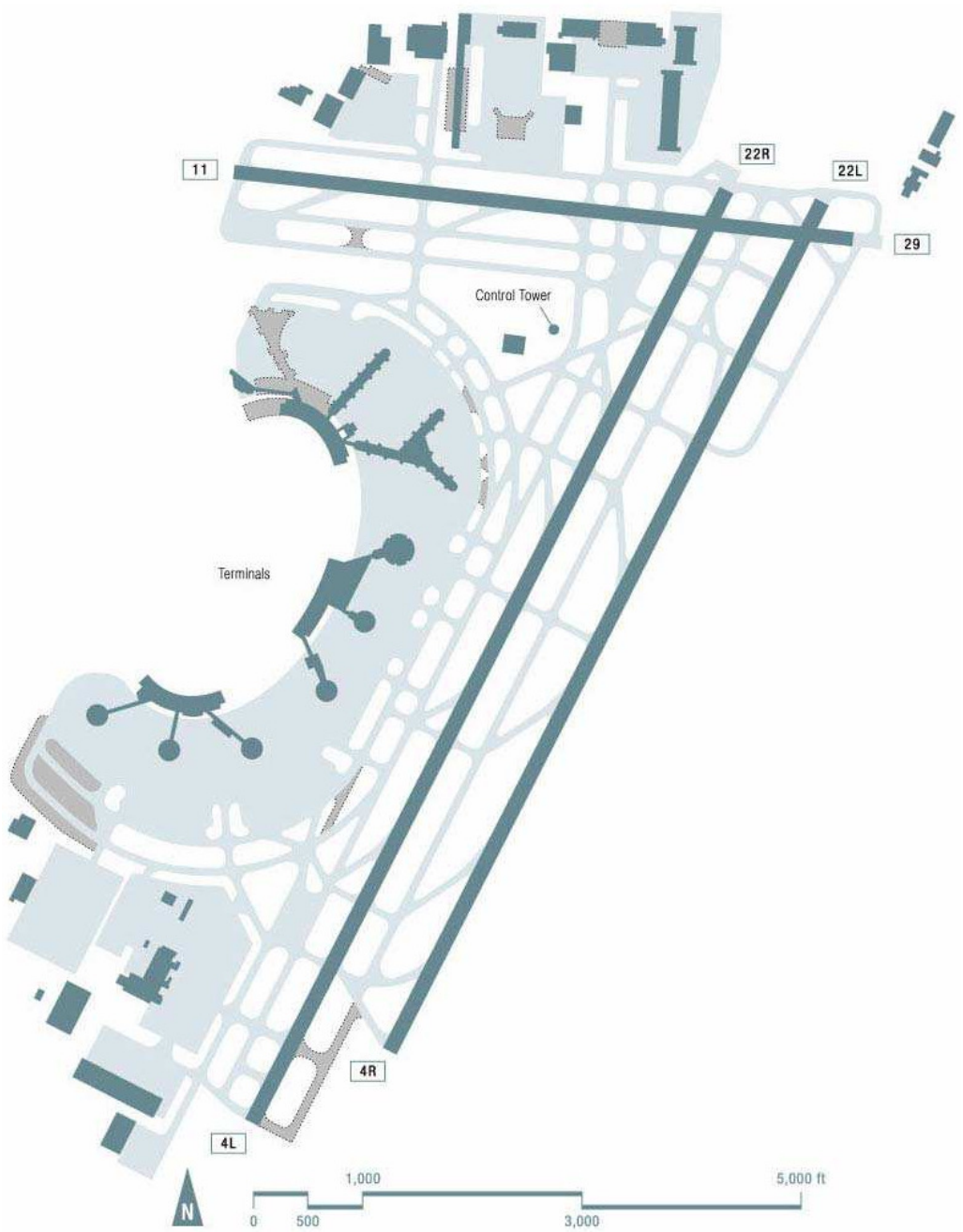


Figure A-2: EWR airport diagram[83]

A.3 La Guardia Airport(LGA)

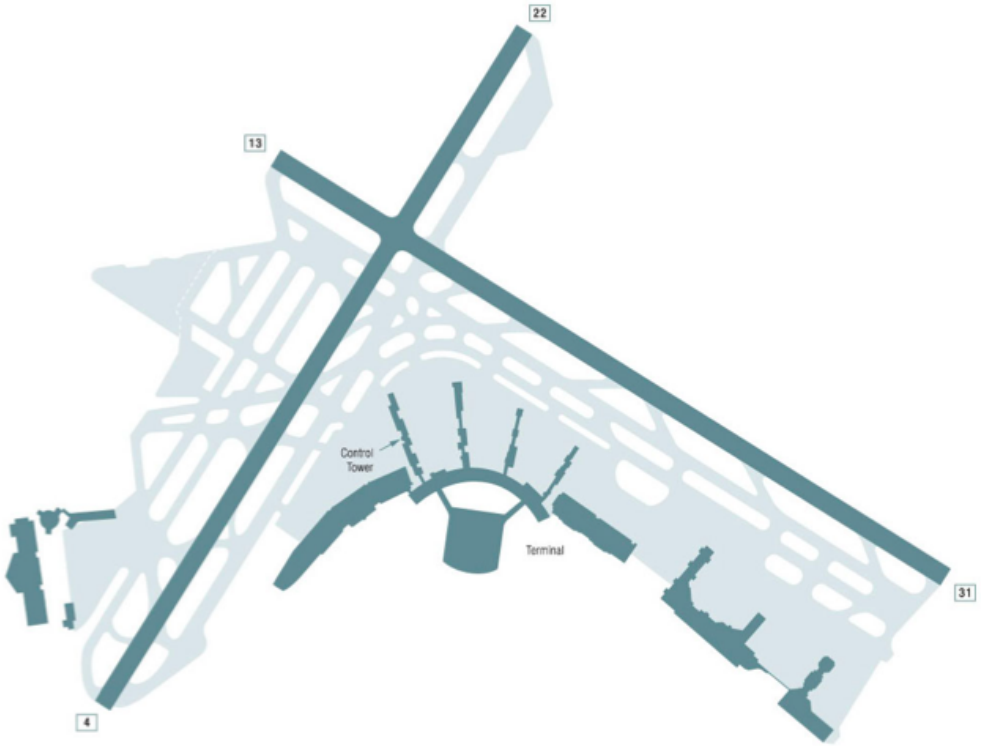


Figure A-3: EWR airport diagram[83]

A.4 Philadelphia International Airport (PHL)

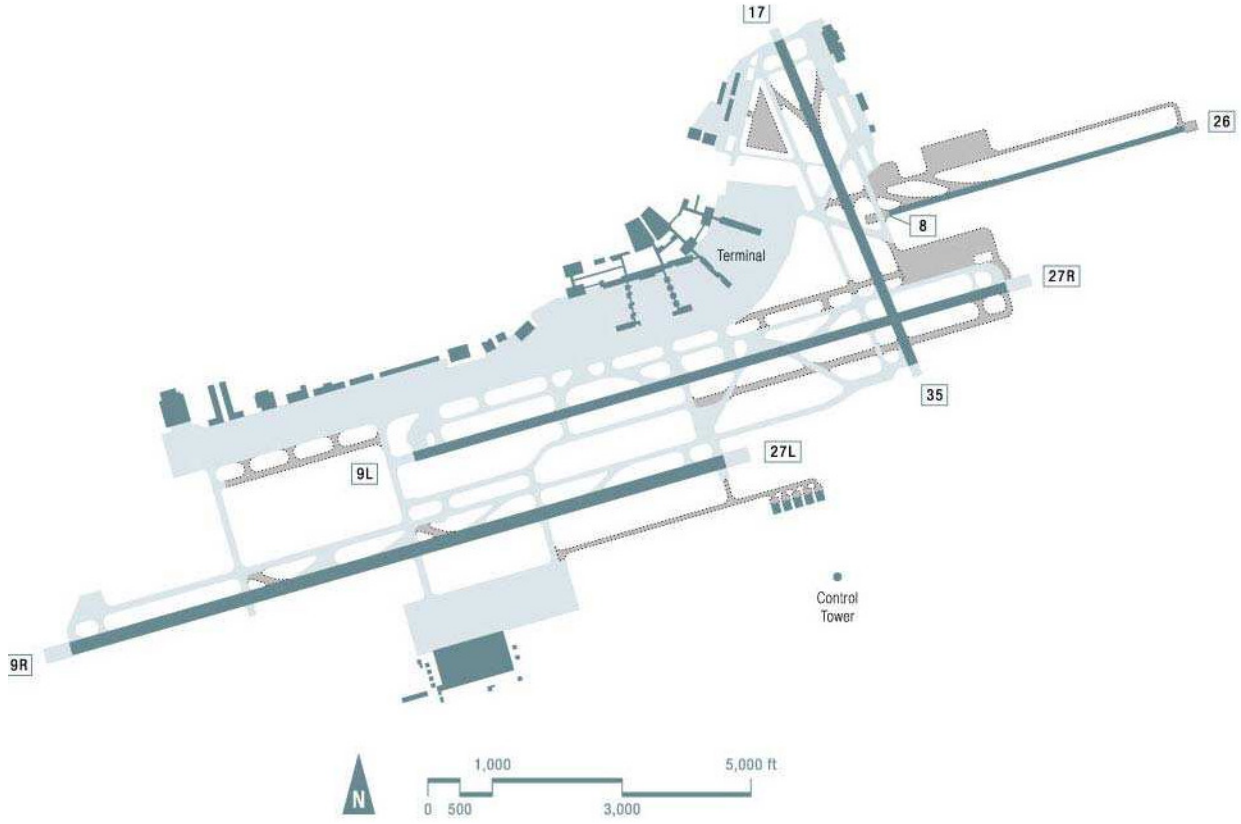


Figure A-4: PHL airport diagram[83]

A.5 Boston Logan International Aiprort(BOS)

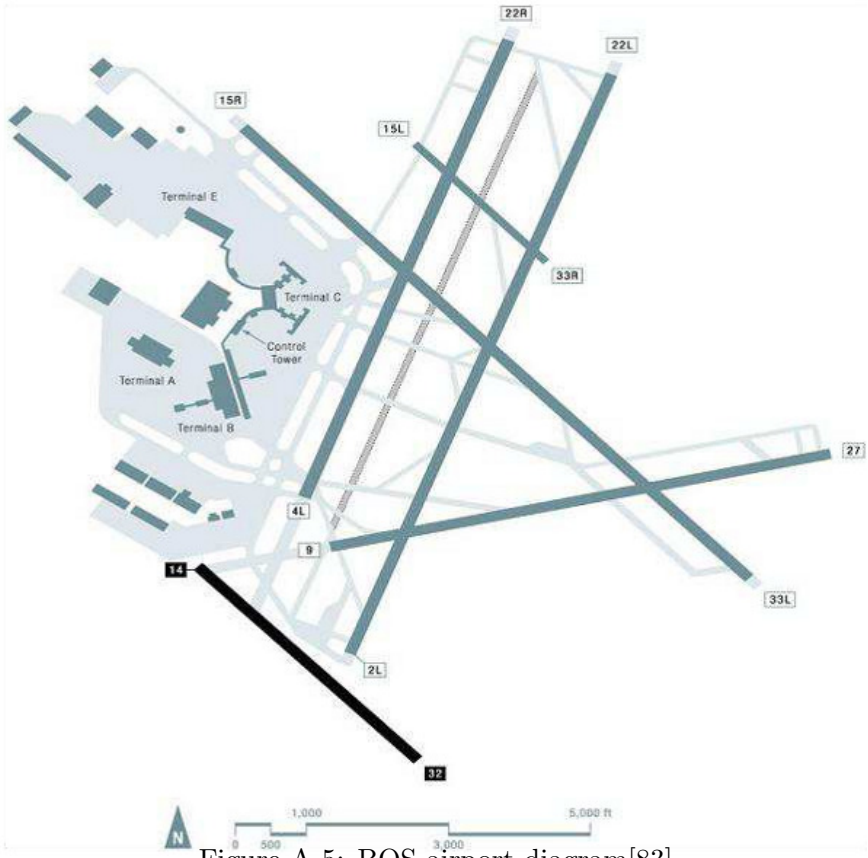


Figure A-5: BOS airport diagram[83]

A.6 Charlotte Douglas International Aiprort(CLT)

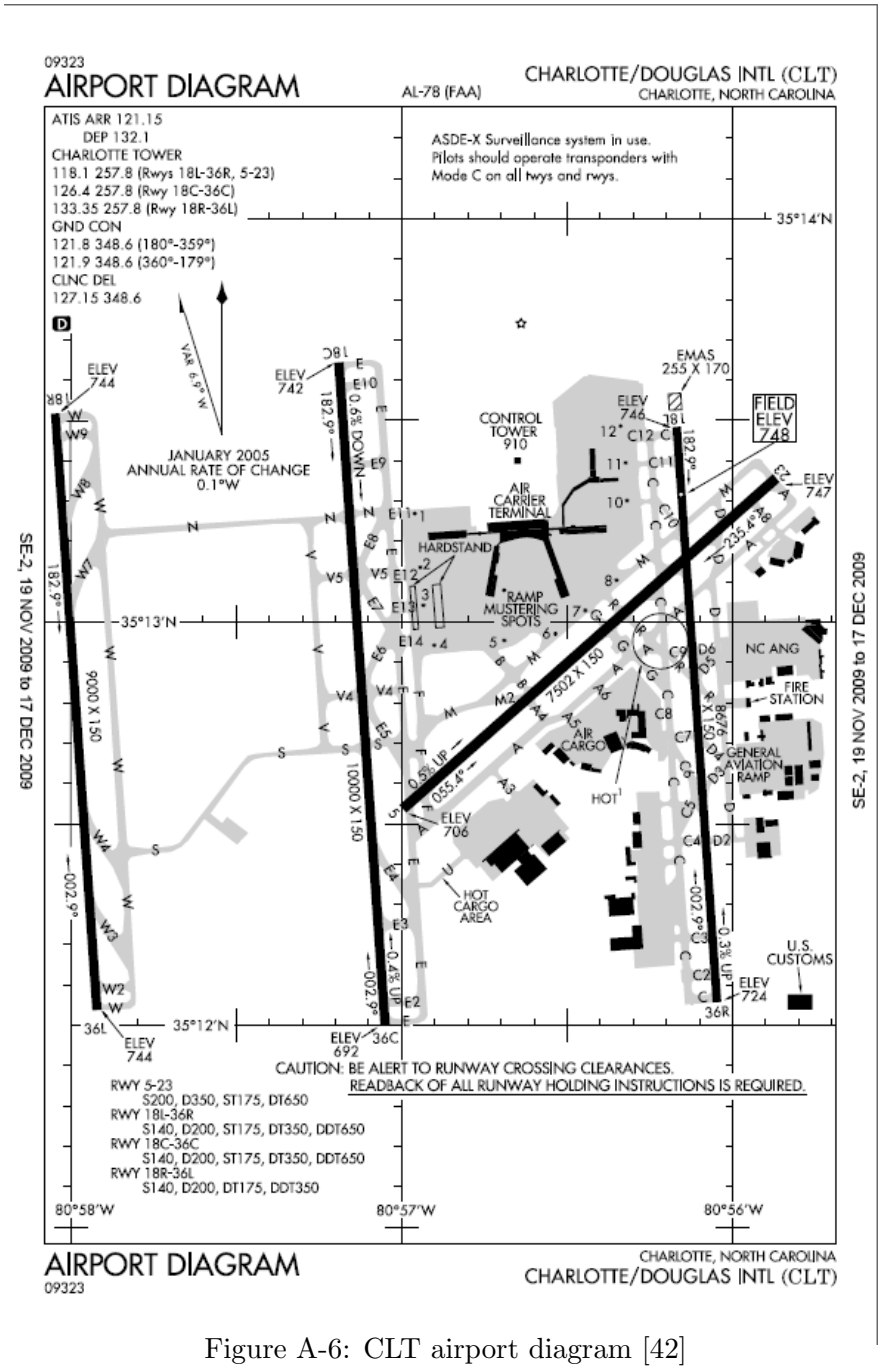


Figure A-6: CLT airport diagram [42]

A.7 Dallas/Fort Worth International Airport (DFW)

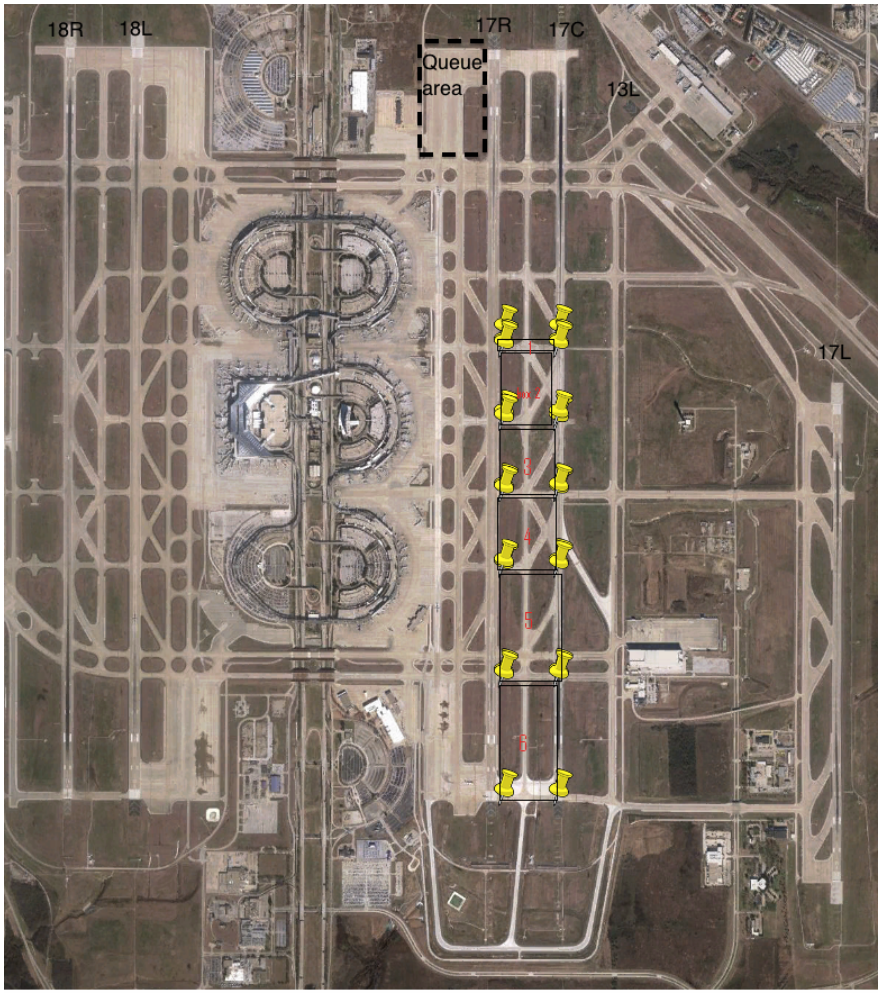


Figure A-7: DFW airport diagram with runway crossing boxes (courtesy of Lincoln Labs)

Appendix B

New York Metroplex Airspace and Airfields

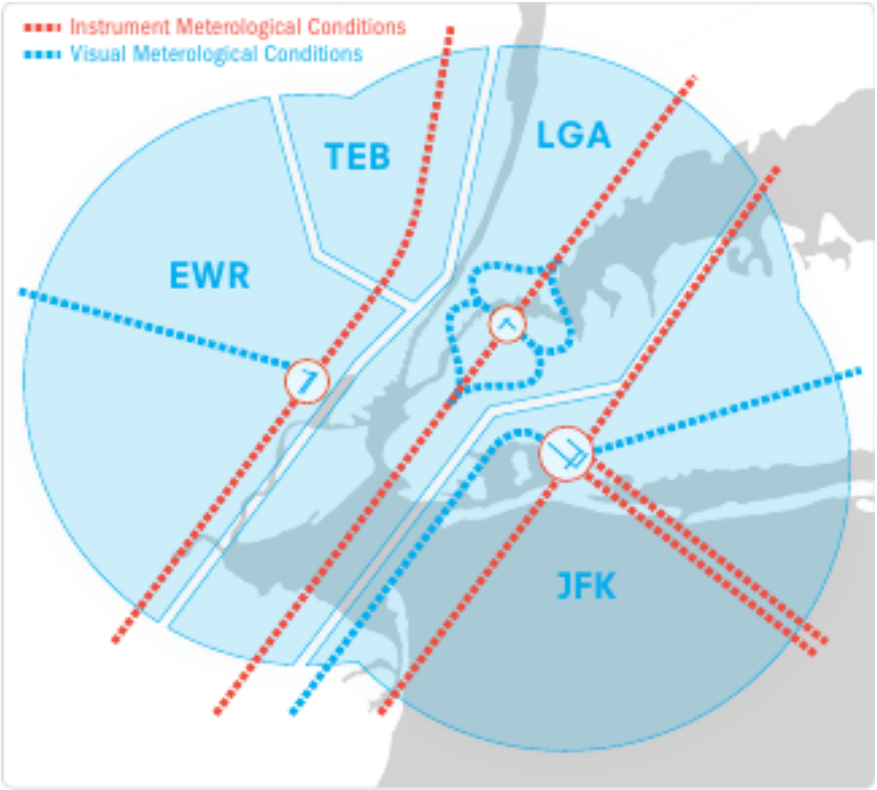


Figure B-1: Airspace Configuration [130]

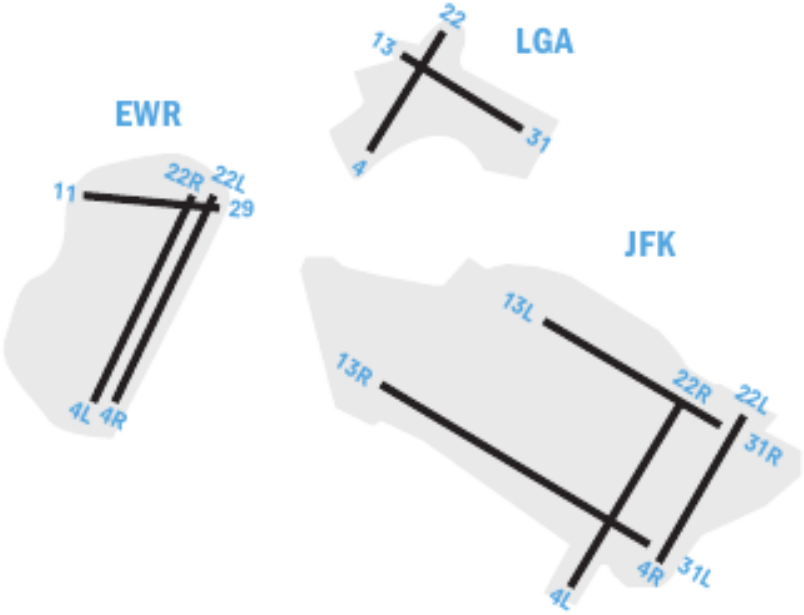
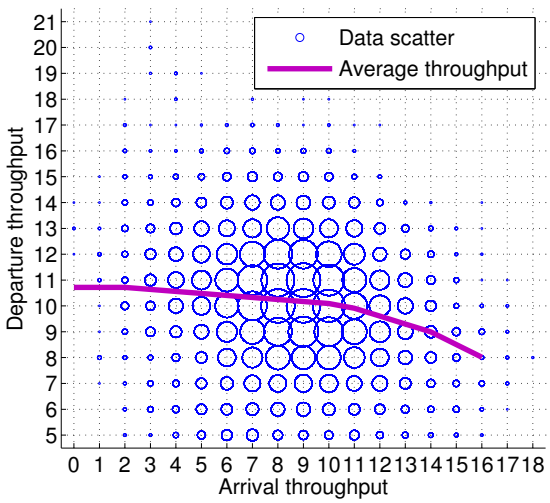


Figure B-2: Relative runways orientation [130]

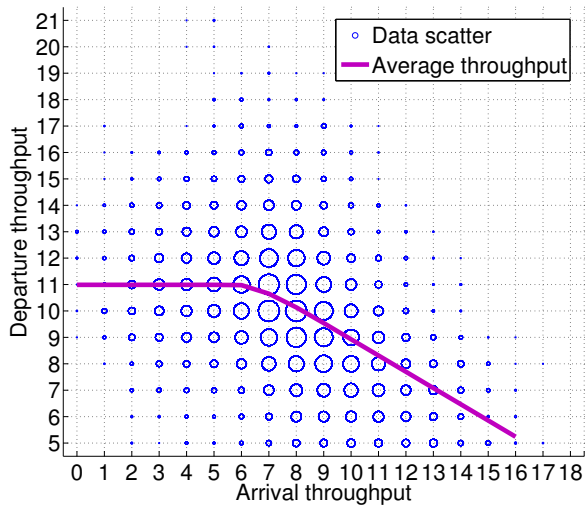
Appendix C

Operational Throughput Envelopes

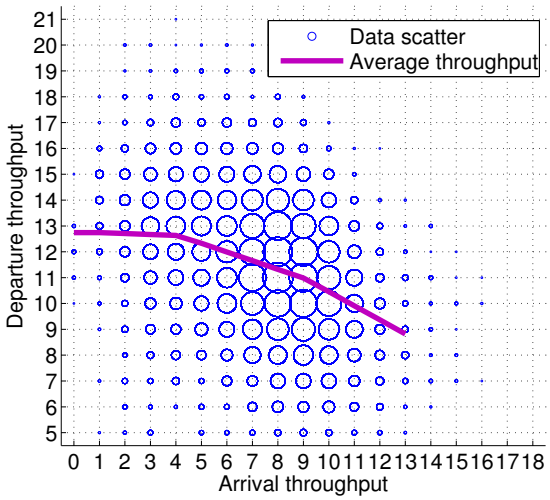
C.1 John F. Kennedy International Airport(JFK)



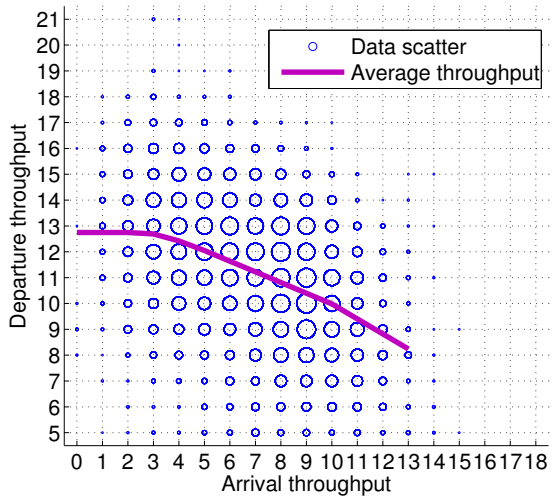
(a) Operational throughput envelope for 13L, 22L | 13R.



(b) Operational throughput envelope for 31L, 31R | 31L.



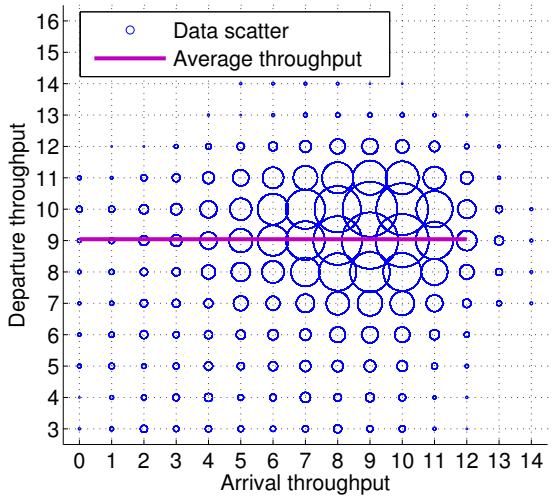
(c) Operational throughput envelope for 22L | 22R, 31L.



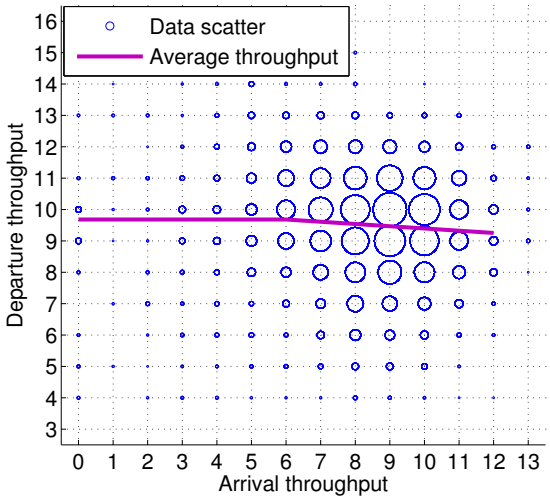
(d) Operational throughput envelope for 4R | 4L, 31L.

Figure C-1: Operational throughput envelopes for the major runway configurations of JFK.

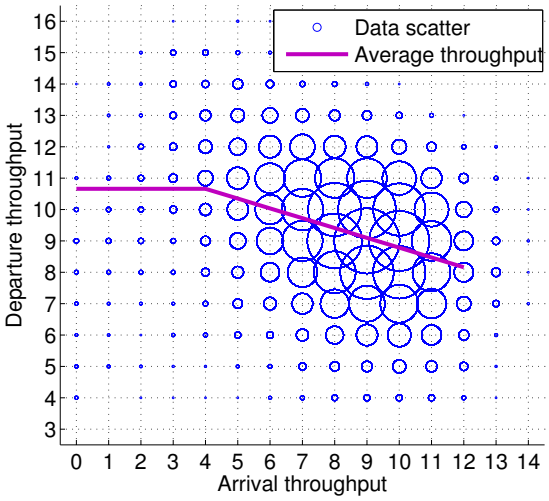
C.2 La Guardia Airport(LGA)



(a) Operational throughput envelope for 22 | 13.



(b) Operational throughput envelope for 22 | 31.



(c) Operational throughput envelope for 31 | 4.

Figure C-2: Operational throughput envelopes for the main runway configurations of LGA.

C.3 Philadelphia International Airport (PHL)

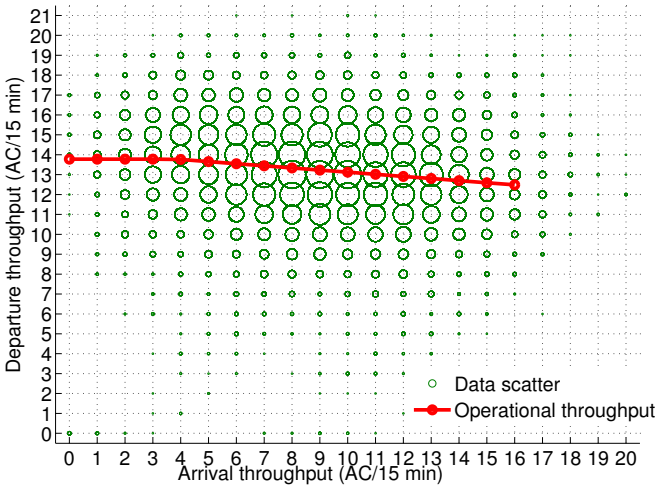
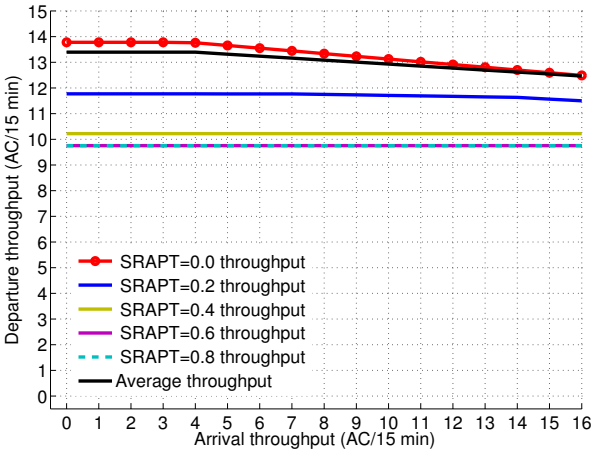
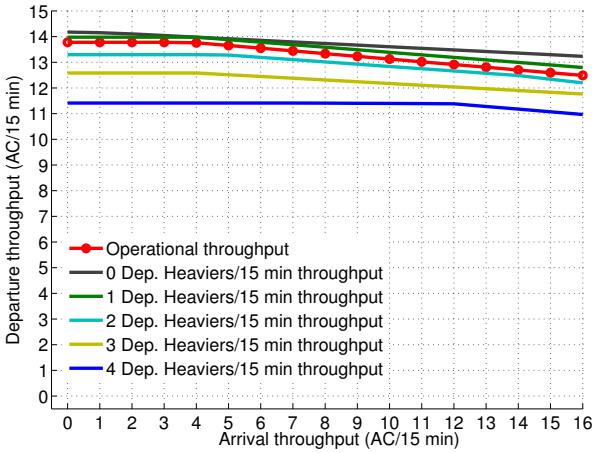


Figure C-3: operational throughput envelope for runway configuration 26, 27R, 35 | 27L, 35 of PHL



(a) Operational throughput envelope parametrized by the SRAPT value



(b) Operational throughput envelope parametrized by the number of Heavy departures

Figure C-4: Parametrized operational throughput envelope for runway configuration 26, 27R, 35 | 27L, 35 of PHL

Appendix D

Analysis of Single Runway Operational Performance: Examples from DFW and PHL

D.1 Introduction

In this chapter, we show how the methodology developed in Chapter 2 can be applied to the capacity analysis of a particular runway using ASDE-X data. First, we perform a detailed estimation of the capacity of Runway 17R at Dallas Fort Worth Airport (DFW) using a small, but detailed, dataset. We show that ASDE-X data validates the conjectures of Chapters 2 and 3 regarding the impact of runway crossings and Heavy departures on departure capacity. Subsequently, we use ASDE-X to measure the departure capacity of Runway 27L at Philadelphia International Airport (PHL), which is not utilized for arrival crossings, and characterize its dependence on Heavy aircraft departures.

D.2 Departure capacity of Runway 17R of DFW

In this section we study the departure capacity of Runway 17R at DFW, using an ASDE-X dataset[1] that includes the times of all the events occurring at the runway. Detailed ASDE-X data, when available, offers the ability to accurately characterize some of the phenomena discussed in this thesis:

[1]The data was provided by MIT Lincoln Labs.

- The differences in inter-departure times resulting from differences in the fleet mix.
- The cost of crossing taxiing traffic through an active runway.

It also addresses some of the limitations of the ASPM data, such as the inability to differentiate between different runways, the requirement for a large dataset containing many congestion periods, and the inherent ambiguity in defining the demand (N) as all aircraft taxiing out.

We present results for the estimation of inter-departure times for Runway 17R of DFW, and the estimation of the capacity of this runway. First, we discuss the departure throughput of the runway. Then, we derive linear regression models and regression trees for the prediction of inter-departure times from a set of selected explanatory variables, such as aircraft type, queue length, arrivals crossing the runway, etc.

D.2.1 Operations at DFW

Table D.1 lists the major runway configurations in DFW. Runway configuration 13R, 17C, 17L, 18R | 17R, 18L, 18R which utilizes four runways for arrivals and two primary runways for departures is used most frequently. The operations of the two departure runways are completely decoupled: Aircraft with filed route departure fixes on the east side of the airport are assigned to east departure runway (17R), and aircraft with filed route departure fixes on the west side of airport are assigned to west departure runway (18L). Thus, the use of the two runways tends to be decoupled and not balanced. In this case, the total number of aircraft taxiing out is not a good predictor of the departure throughput, we also need information about the departure runway of each aircraft to identify if a runway is in saturation. Similarly, the arrival throughput and the interaction of arriving aircraft with departures can differ substantially depending on how the traffic is allocated to the four runways.

Table D.1: Major runway configurations in DFW under Visual Meteorological Conditions.

Name	Arrival Runways	Departure Runways	Frequency of use
South–flow	13R, 17C, 17L, 18R	17R, 18L, 13L (props)	60%
North–flow	31R, 35C, 35R, 36L	31L (props), 35L, 36R	39%
North–west	31L, 31R	31L, 31R	1%

Figure D-1 reveals these challenges. We notice that the saturation throughput exhibits an unstable behavior. It nears 20 operations/ 15 minutes when departure demand is between 20-25 aircraft, subsequently drops and increases again at higher values of departure demand. This cyclic

behavior could be attributed to different utilization of the three departure runways at different high-demand periods. An overview of the imbalanced use of the departure runways and its impact on taxi-out related delays in DFW was previously presented by Atkins and Walton [4].

It is not clear how to define departure throughput and to derive the filtered dataset in saturation without knowing the demand at each individual runway. We show in the remainder of this section how to measure the departure throughput of an individual runway of this runway configuration (17R), and identify the explanatory variables that explain some of its variation using 11 days of ASDE-X data from 2009.

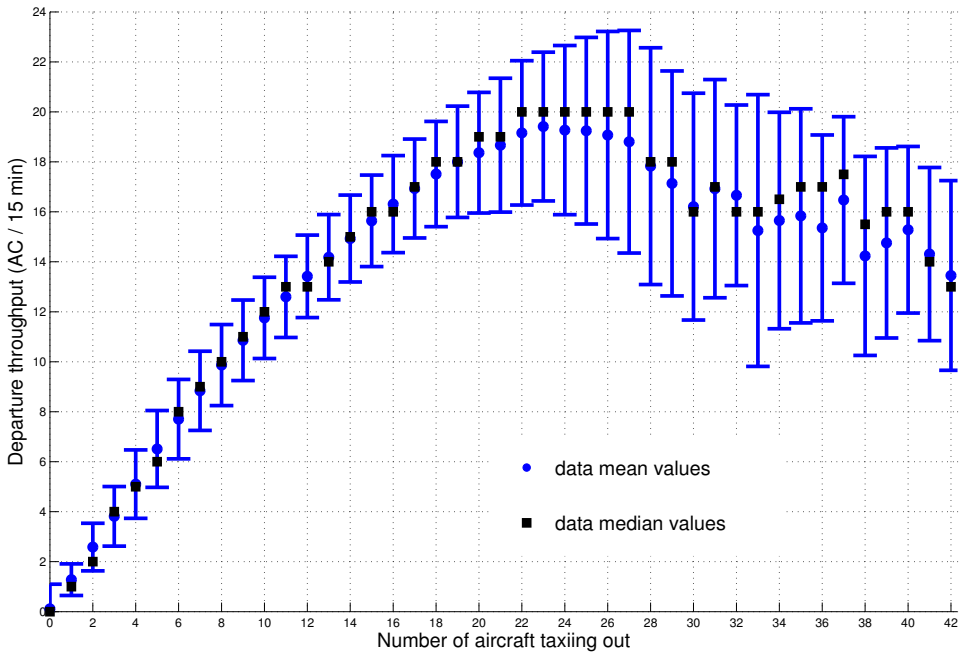


Figure D-1: Departure throughput as function of the number of aircraft taxiing out in 2010.

D.2.2 Departure throughput as a function of the departure queue

First, we construct the saturation throughput plot for Runway 17R alone, instead of the airport as a whole. For this, we measure the departure throughput of Runway 17R using ASDE-X data. The second step is to measure the departure demand for the runway accurately. For this, we measure the number of aircraft in the queue area, shown in Figure A-7. This defines the departure demand for this runway. This queue area is 750 meter long and wide enough to fit three aircraft abreast, offering a queuing capacity of more than 30 aircraft. This queue area is further divided into three parallel sub-queues [67].

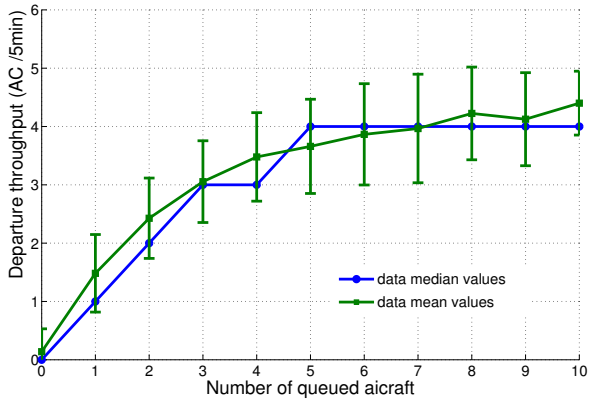
The departure throughput of Runway 17R is represented as a function of the departure demand, $Q(t)$, measured as the number of aircraft in the departure queue of Runway 17R during minute t. In other words, it is the number of taxiing out aircraft that have reached the departure queue, but not taken off yet.

$\bar{T}_{dt}$ is defined as the number of takeoffs over the time period $[t, t+1, ..., t+dt)$. Adapting the method suggested by Pujet [91], we find the time interval dt for which $Q(t)$ and $\bar{T}_{dt}$ have the highest correlation. In this case, the highest correlation between $Q(t)$ and $\bar{T}_{dt}$ is obtained for $dt = 5$, implying that the number of aircraft in the departure queue at time t, namely $Q(t)$, is a good predictor of the number of takeoffs during the 5-min time interval $[t, t+5)$. This observation reflects the lack of significant congestion in DFW. Since the capacity of the queue area is higher than 30 aircraft and yet the number of aircraft in queue predicts the throughput at best only for a 5-min horizon, the aircraft in queue will depart on average within the next 5 minutes.

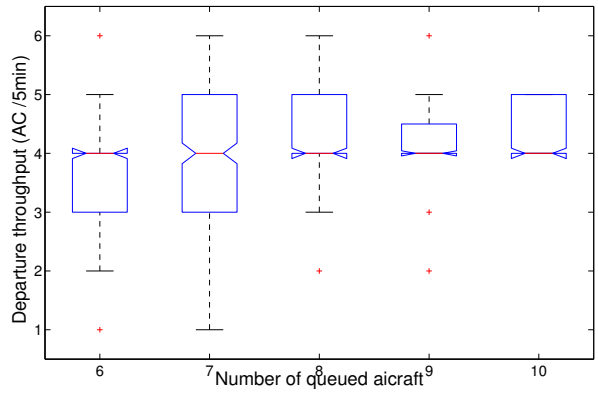
This representation yields the plots of Figure D-2a for the 11 days of 2009. In total, we have 15,840 data points (Q, $\bar{T}_{dt}$)-one for every minute. The mean and median values of the takeoff rate for each value of the departure demand, Q, are plotted. The error bars depict the standard deviation of the takeoff rate at each value of Q.

From Figure D-2a, one observes that the departure throughput of this runway saturates at around 4 takeoffs/ 5 min when the queue is around 5 aircraft or longer. A robust way to identify the queue length beyond which the throughput does not vary with the queue length is to use the algorithm presented in Section 4.4.4. We group the throughput observations of each value of Q, and use the Kruskal-Wallis one-way analysis of variance test to test for significant differences between the empirical throughput distributions of each group. The test does not reject the null-hypothesis that the throughput observations at different values of Q are drawn from the same distribution at both 0.05 and 0.1 significance levels for $Q \geq 6$. However, it does reject the null hypothesis if more groups at lower values of Q are included.

The test implies that the measurements of throughput for different values of Q when $Q \geq 6$ are not significantly different. We define the inter-departure time between two aircraft under persistent demand (that is, when the departure queue has six or more aircraft) as the service time of a departure. The average service time is around 75 sec. In the rest of the analysis, we examine the parametrization of this estimate using more information about fleet mix, arrivals crossings and the departure queue.



(a) Departure throughput as function of departure demand.



(b) Boxplots of the departure throughput for high values of departure demand.

Figure D-2: Departure throughput of Runway 17R as function of number of aircraft in the departure queue.

Table D.2: Service time as a function of queue size.

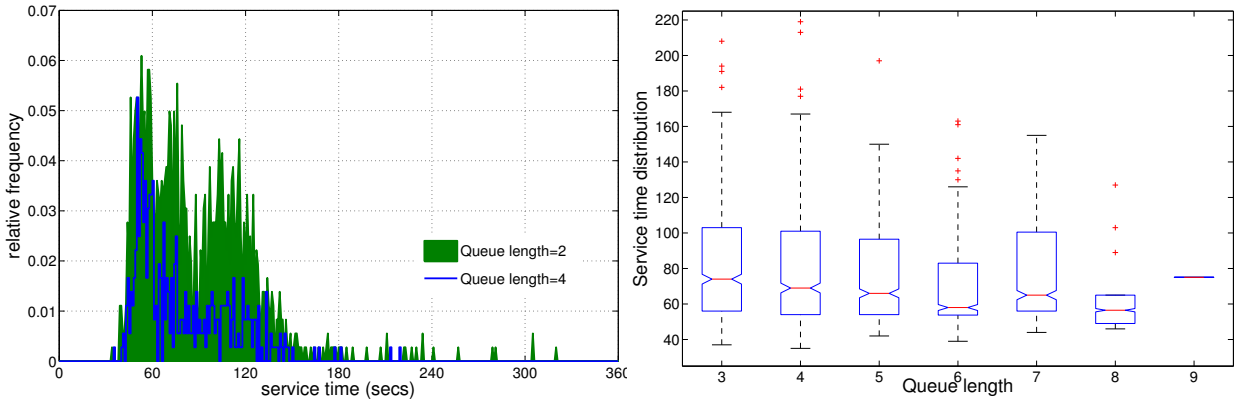
Queue Length (AC)	Mean Service time (sec)	Median service time (sec)	Number of observations
1	104	97	1,324
2	89	81	1,102
3	81	74	719
4	79	69	421
5	78	66	185
6	73	58	85
7	80	65	31
8	65	56	14
9	75	75	1

D.2.3 The service time as a function of the queue length

As was done in Section 4.4.4, we define the queue length ($d_q(i)$) for each aircraft i as the number of aircraft in the departure queue at the moment when aircraft i starts its takeoff roll. We filter out all flights that started their takeoff roll when the departure queue was empty. For all the remaining flights, some other flight was in the departure queue when they took off. We first study how the service time changes with the queue length. In Table D.2, the sample mean and median service times as a function of the queue length at the start of the takeoff roll of a departing aircraft are presented. One can observe that both the mean and the median service time seem to decrease as the number of queued aircraft increases from 1 to 3. Beyond 3, the average service time fluctuates around between 60 and 80 sec. We use the Kruskal-Wallis test to verify this inspection.

We group the throughput observations of each unique value of d_q, and use the Kruskal-Wallis test for significant differences between the service times of each group. The test does not reject the

null-hypothesis that the service times for different values of d_q are drawn from the same distribution at both 0.05 and 0.1 significance levels for $d_q \geq 3$. However, it does reject the null hypothesis if we include more groups at lower values of d_q. The empirical distributions of the service time for two different values of queue length (2 and 4) are visualized in Figure D-3a. The distinction between the two distributions is clear. In Figure D-3b, we depict the boxplot of the service times for queue length 3 and greater. One can observe the similarity between the boxplots.



(a) Empirical distribution of the service time for queue length 2 and 4.

(b) Boxplots of the service time as a function of the queue length.

Figure D-3: Service time distributions.

In the subsequent analysis, we model and predict the service time. For this, we filter out all flights for which the queue length was less than 3 when they started their takeoff roll. When the queue length is smaller than 3, the inter-departure time between two flights might be longer, because the trailing flight may not be at the runway threshold at the end of the separation requirement. We also note that at DFW, one of the sub-queues is occasionally used as a holding area.

D.2.4 A linear regression model for the inter-departure time prediction

We formulate a linear regression model for the prediction of the service time of a departing aircraft, conditioned on the queue length being greater than or equal to 3. We want to predict the inter-departure times as a function of the aircraft type, the arrivals crossings and the departure queue. In this section we describe the potential explanatory variables considered.

The inter-departure time is expected to be around 2 min if the leading aircraft is Heavy or B757. Otherwise, it is expected to be 1 min or less. If the leading and trailing aircraft are assigned to different departure fixes, the inter-departure time is expected to equal the required time for the leading aircraft to clear the departure runway. For an average jet, this time is around 45 sec. If the

leading and trailing non-Heavy aircraft head to the same departure fix, 2.5 miles spacing between the two must be ensured which translates to approximately 60 sec inter-departure time.

In this runway configuration, aircraft arriving on runways 17C and 17L need to cross the departure Runway 17R in order to reach the terminal area. These runway crossings could impact the availability of the runway for departures, as recognized in other studies [69]. It is also hypothesized that the further from the runway threshold that the crossing occurs, the higher the cost to runway utilization will be. Taxiing-in aircraft can cross the runway as soon as the leading departure has passed the crossing point. If the crossing is close enough to the intersection, the taxiing-in aircraft has a high probability of crossing the runway within the separation requirement between the two departures. In this case, it has no impact on the departure throughput. Similarly, two taxiing-in aircraft crossing the departure runway in a staggered fashion, that is, through parallel intersections, are hypothesized to occupy the departure runway for less time than two crossings through the same intersection, because the staggered crossings can occur almost simultaneously. The different taxiways that cross Runway 17R and are used in the dataset are shown in Figure A-7.

Finally, the queue length could be an explanatory variable, since longer queues may impact the controllers' efficiency. If the next departure is selected from the same subqueue or not, could also be an explanatory variable. This could be the case, if aircraft are organized into sub-queues according to their departure fix. In this case, if the leading and trailing aircraft are selected from a different subqueue, they will have a shorter inter-departure time.

In summary, we consider the following potential explanatory variables in order to model the above mentioned dependencies:

1. *Heavy*: Binary variable, equals 1 if the departing aircraft is Heavy, 0 otherwise.

2. *B757*: Binary variable, equals 1 if the departing aircraft is some type of B757, 0 otherwise.

3. *Small*: Binary variable, equals 1 if the departing aircraft is small, 0 otherwise.

4. *Queue length*: Discrete variable, equals the queue length, greater than or equal to 3.

5. *Same subqueue*: Binary variable, equals 1 if the next departure is selected from the same subqueue, 0 otherwise.

6. *Arrivals crossing*: Binary variable, equals 1 if some arriving aircraft crosses Runway 17R after the departing aircraft takes off, 0 otherwise.

7. *First box*: Discrete variable, equals zero if *Arrivals crossing* is zero. Otherwise: the box-number of the first runway crossing.

8. *Last box*: Discrete variable, equals zero if Arrivals crossing is zero. Otherwise: the maximum box-number of any runway crossing during the service time.

9. *Maximum non-staggered crossings*: Discrete variable, equals zero if *Arrivals crossing* is zero. Otherwise: the largest number of aircraft that use the same box to cross Runway 17R minus one.

10. *Additional crossings*: Discrete variables, equals zero if *Arrivals crossing* is zero. Otherwise: the total number of crossings minus (*Maximum non staggered crossings*) minus one.

In case of a runway crossing, a set of four explanatory variables is used to evaluate the impact of the crossing(s). We measure the total number of crossings that take place during the inter-departure time. The *first box* variable takes the value of the box that the first crossing used. If there are multiple runway crossings, we have: The *last box* takes the value of the maximum box-number used by any arrival to cross the runway. For the reasons explained above, we separate the additional crossings during the inter-departure time in two variables: The *Maximum non-staggered crossings* takes the value of the crossings that use the same box minus one. We subtract one because the first crossing is already accounted. The *Additional crossings* variable equals the number of the remaining runway crossings.

We consider modeling the service time using two different types of linear models:

- Multiple linear regression.
- Regression trees.

For the case of the multiple linear regression, we use the LASSO method for variable selection. The regression trees are pruned using 10-fold cross-validation.

The tree models are very useful for the investigation of complex interactions between the primary explanatory variables. In a tree model, as depicted here, the longer the branches, the greater the deviance explained. The value at the end leaf of each branch is the conditional mean of the taxi out time under the conditions that the branches dictate. Figure D-4 indicates that the *Arrivals crossing* variable (i.e., whether there is a runway crossing from some arriving aircraft) is by large the most critical factor affecting inter-departure times.

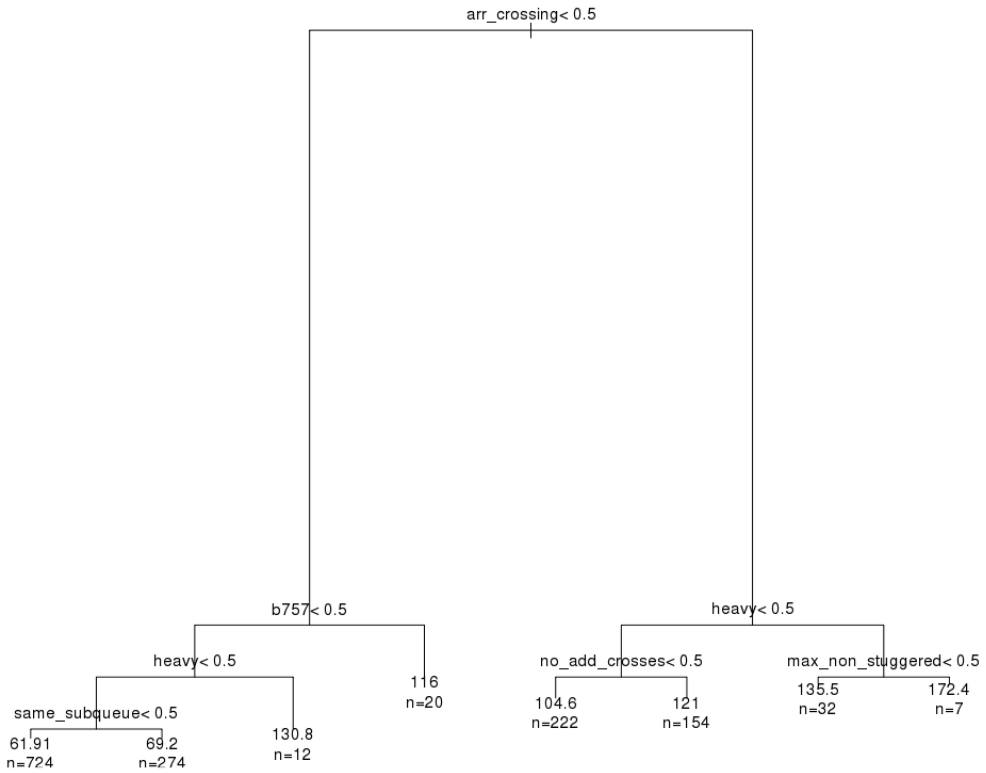


Figure D-4: Regression tree for the service time prediction.

For this reason, we construct two different linear regression models: The first one for the case that *Arrivals crossing*=0 and the second one for the case that *Arrivals crossing*=1.

Linear regression model given zero arrivals crossings

In this case, we only need to choose between the first 5 potential explanatory variables. We find that all of them except for the variables *Small* and *Queue length* have some statistical significance.

Table D.3 lists the resulting coefficients of the linear regression model. The "default" separation time is around 60 sec, and both a Heavy and a B757 introduce longer spacing, as required by regulations. Interestingly, the B757 introduces a 16 sec shorter spacing than a Heavy. We also notice that the coefficient of the *Same subqueue* variable is non-zero. This suggests that the event of the next plane being selected from the same subqueue introduces a longer inter-departure time, as hypothesized.

We note from Table D.4 that the significance of the regression is small. The R^2 value is 0.28. This was expected since we have conditioned the linear regression on no arriving aircraft crossing the departure runway and the queue being sufficiently large (≥ 3). Given these conditions and the homogeneity of the fleet mix in DFW, it was expected that the remaining explanatory variables

Table D.3: Intercept and regression coefficients for the multiple regression model given zero arrivals crossings.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value	p-value
(Intercept)	62	1	99	$< 2 \cdot 10^{-16}$
Same subqueue	7	1	6	$9 \cdot 10^{-09}$
Heavy	67	5	14	$< 2 \cdot 10^{-16}$
B757	51	4	13	$< 2 \cdot 10^{-16}$

Table D.4: Summarized results for the multiple regression model given zero arrivals crossings.

SER	16.94 on 1026 degrees of freedom
R^2	Multiple R^2: 0.2798 Adjusted R^2: 0.2777
F-statistic	132.9 on 3 and 1026 DF, p-value: $< 2.2e - 16$

Table D.5: Intercept and regression coefficients for the robust regression model given zero arrivals crossings.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value
(Intercept)	60	1	115
Same subqueue	6	1	6
Heavy	69	4	17
B757	53	3	16

would not explain much of the variance of the inter-departure times. Indeed, the sample mean is 66 sec and the sample standard deviation 20 sec. The regression model achieves a more accurate estimate than the sample mean for the few cases that the next departure is selected from the same subqueue (less than 28% of the time), and the even fewer times that the leading departure is a Heavy or a B757 (less than 4% of the time).

The difference between the mean and median service time, shown in Table D.2 and the fat tail of the distribution of the service times, shown in Figure D-3a suggest that the regression coefficients could be biased by the presence of outliers. Thus, we also consider a more robust version of the linear regression model, namely M-estimators (using Huber's ψ function with the default turning constant of 1.345). The resulting coefficients are presented in Table D.5. Comparing the results listed in Tables D.4 and D.5, we note that all three explanatory variables remain significant and that their estimated values do not change substantially.

Because of the simplicity of the model and the binary nature of most explanatory variables, the tree model looks very similar (Figure D-5). It additionally suggests that the queue length could also affect the inter-departure time: When it is longer than 3, the inter-service time is on average 6 sec shorter under certain conditions.

One could simplify this model even further by removing the *Same subqueue* variable. This is useful if the value of this variable cannot be known in advance. We list the coefficients and the

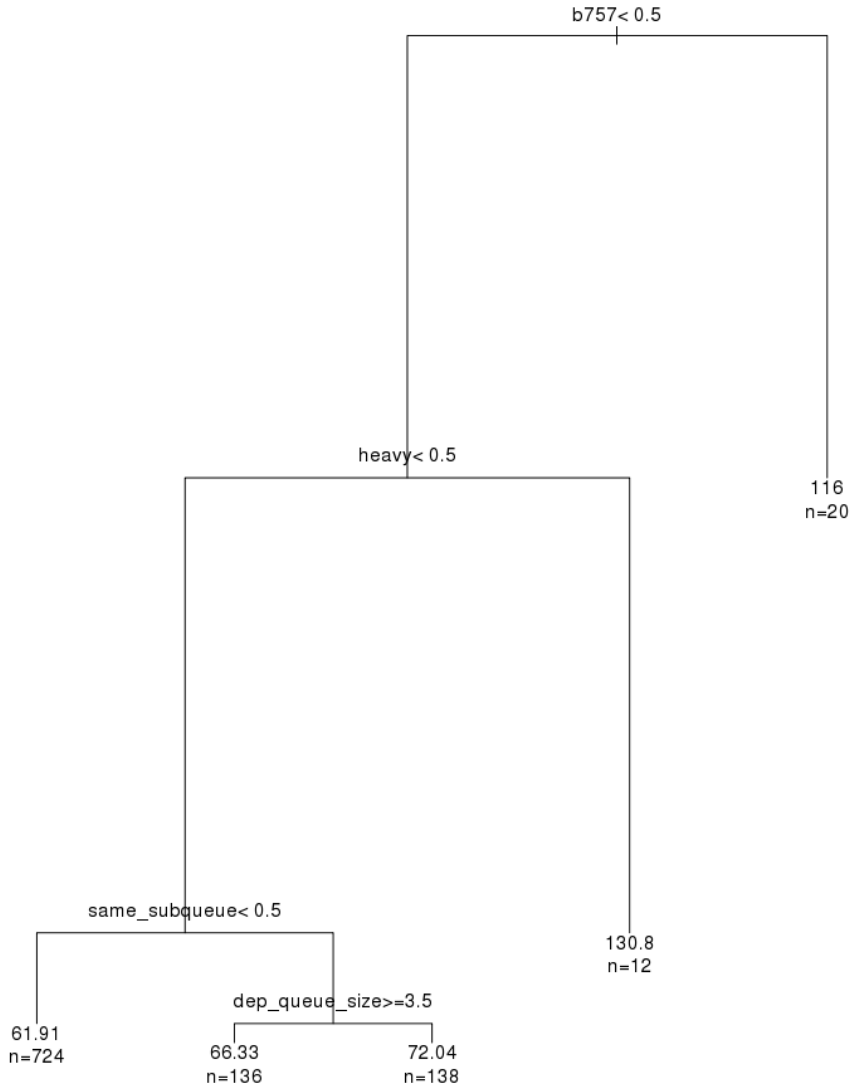


Figure D-5: Regression tree for the service time prediction given zero runway crossings.

results of this regression model in Table D.6 and Table D.7. The M-estimators are listed in Table D.8. The estimated values for the intercept and the coefficients do not vary significantly with the removal of the *Same subqueue* variable. This suggests that the estimated coefficients are not subject to multicollinearity, and that they accurately measure the impact of each explanatory variable.

Linear regression model given nonzero arrivals crossings

For the regression model given nonzero arrivals crossings, the variable selection process is more challenging since we have to select the most significant explanatory variables out of nine potential explanatory variables (The sixth variable, *Arrivals crossing*, equals 1). We use again the LASSO method for variable selection. We list the selected variables, their coefficients, and the results of the

Table D.6: Intercept and regression coefficients for the multiple regression model given zero arrivals crossings without the "Same subqueue" variable.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value	p-value
(Intercept)	64	1	117	$< 2 \cdot 10^{-16}$
Heavy	67	5	13	$< 2 \cdot 10^{-16}$
B757	52	4	13	$< 2 \cdot 10^{-16}$

Table D.7: Summarized results for the multiple regression model given zero arrivals crossings without the *Same subqueue* variable.

SER	17 on 1027 degrees of freedom
R^2	Multiple R^2: 0.26 Adjusted R^2: 0.26
F-statistic	177 on 2 and 1027 DF, p-value: $< 2 \cdot 10^{-16}$

regression model in Table D.9 and Table D.10. The M-estimators are listed in Table D.11. Similarly to the case of zero arrivals crossings, the M-estimators are very close to the linear regression estimated coefficients. In fact, for 4 out of 5 estimators, they are within the standard error of the estimated value.

Comparing Table D.3 and Table D.9, we note the remarkable difference between the values of the intercept. If all the other variables equal zero, the mean inter-departure time increases from 62 to 104 sec because of a single arriving aircraft crossing Runway 17R. This implies that the cost of a runway crossing is larger than 40 sec. However, this cost gets partially offset by the significant difference in the coefficients of the variables Heavy and B757, which both decrease for about 40 sec. This results in approximately same service time of a Heavy or a B75 (62+67 sec and 104+30 sec for a Heavy and 62+51 sec and 104+12 sec for the B757) independently of the value of *Arrivals crossing*. After a Heavy or B757 takeoff, the controllers utilize the long wake vortex separation requirement (120 sec) to perform a runway crossing at almost no additional cost. In fact, the service time of a departure of a Heavy or a B757 and an arrival crossing increases on average for less than 5 sec compared with the service time of just a departure of a Heavy or a B757.

This finding also provides further evidence supporting the conclusions for the impact of Heavy

Table D.8: Intercept and regression coefficients for the robust regression model given zero arrivals crossings without the *Same subqueue* variable.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value
(Intercept)	62	0	137
Heavy	69	4	17
B757	54	3	17

Table D.9: Intercept and regression coefficients for the multiple regression model given nonzero arrivals crossing.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value	p-value
(Intercept)	104	1	81	$< 2 \cdot 10^{-16}$
Heavy	30	3	9	$< 2 \cdot 10^{-16}$
B757	12	3	4	$4.145 \cdot 10^{-5}$
Additional crossings	7	1	5	$4 \cdot 10^{-7}$
Maximum non staggered crossings	11	3	3	$2 \cdot 10^{-3}$

Table D.10: Summarized results for the multiple regression model given nonzero arrivals crossing.

SER	20 on 410 degrees of freedom
R^2	Multiple R^2: 0.32 Adjusted R^2: 0.31
F-statistic	48 on 4 and 410 DF, p-value: $< 2 \cdot 10^{-16}$

Table D.11: Intercept and regression coefficients for the robust regression model given nonzero arrivals crossing.

Coefficient	Estimate (sec)	Std. Error (sec)	t-value
(Intercept)	103	1	101
Heavy	26	3	9
B757	12	2	4
Additional crossings	7	1	6
Maximum non staggered crossings	10	3	4

and B757 aircraft departures in the case of BOS runway configuration 22L, 27 | 22R, 22L, which is similar to this runway configuration of DFW. The primary departure runway is 22L and arrivals on runways 22L and 27 have it cross it to reach the terminals. The curves of Figure 2-14b show that under medium or heavy arrival rates, the 15-min departure throughput does not change for 0-4 Heavy or B757 departures. Medium or heavy arrival rates require many runway crossings. In DFW, if 2 crossings are performed after a Heavy or B757 departure, the total runway time utilized for runway crossings is the same as performing each crossing after a Large aircraft departure. With such a strategy, the cost of a Heavy or B757 departure diminishes.

The significant difference in the coefficients of the intercept and the variables Heavy and B757 in Tables D.3 and D.9 also shows why it was necessary to construct two separate regression models conditioned on the value of the *Arrivals crossing* variable.

Comparing Table D.4 and Table D.10, one can note that the standard error of the regression (SER) increases for the model of nonzero arrivals crossing. However, the R^2 coefficient increases because the explanatory variables explain more of the variability of the data than for the case of zero arrivals crossing. Given *Arrivals crossing=1*, the mean service time is 114 sec and the standard deviation 24 sec. The linear regression model of Table D.10 explains 31% of this variability.

Concerning the selection of variables, it turns out that the box, or boxes that are used for the crossings are not significant explanatory variables in the linear regression model. However, additional crossings lead to longer service times. As hypothesized, the non-staggered crossings introduce a longer delay than the staggered ones. However, the difference is not significant. For both types of crossing groupings, the delay for an additional crossing is much shorter than the delay of the first crossing (7 and 10 vs. 42 sec). This suggests that grouping the crossings is more efficient, even if they are routed through the same box. The marginal cost of any additional crossing is at most 10 sec. However, the cost of the first crossing is larger than 42 sec.

D.2.5 Operational throughput envelope of Runway 17R

The statistical analysis of Section D.2.4 is important for estimating the dependency of the inter-departure time on several explanatory variables, but does not help us characterize the operational capacity of the runway, as explained in Section 4.4.4. For example, in the presence of a demand for five Large departures and five runway crossings, the throughput (and the delays) will be very different in the following two strategies:

1. The controllers choose to alternate crossings and departures.

2. The controllers perform all crossings staggered together after the last departure.

According to the results of Sections D.2.4 and D.2.4, the total time required would be 5×104 sec $= 520$ sec in the first case, which implies a throughput of 35 departures and 35 runway crossings/hr. In the second case, the time required would be $64 \times 4 + 104 + 7 \times 4$ sec $= 388$ sec. This strategy would imply a throughput of 46 departures and 46 runway crossings/hr. The throughput of the second strategy is 31% higher than the throughput of the first strategy.

The aggregate estimation methods proposed in this work, like the operational throughput envelope, address these issues by estimating the average number of departures and arrivals of the airport over a longer time period, usually 15 minutes. In this case, we estimate the performance only over a 5 minute period, because of data scarcity: Over the 11 days of ASDE-X data, we do not have queue loads that would allow us to observe the departure throughput under high-demand over a 15-minute period. As we saw in Section D.2.3, the longest queue is 9 aircraft and is observed only once.

We extend the framework of Chapter 2 for estimating and representing the departure capacity of Runway 17R as a function of the taxiing-in aircraft crossing the runway, and the number of Heavy aircraft and B757s in the departing fleet mix. As established in Section D.2.2, the saturation condition for this case is having 6 or more aircraft in queue. There are only 550 datapoints that satisfy this condition. Under this condition, the departure throughput in the next 5-minute interval does not change with the number of aircraft in the queue box. The scatter plot and the fitted function for the departure throughput for different numbers of arrivals crossings is shown in Figure D-6a.

The formulation to determine the fitted function is similar to the one of Section 2.3, but simpler. We define $\bar{T}(t)$ as the number of aircraft that take off during the 5-minute interval $[t, t+5)$ min. Similarly, we define $\bar{Z}(t)$ as the number of runway crossings that are conducted during the 5-minute interval $[t, t+5)$ min. Given k pairs of measurements $\bar{Z}(t)$ and $\bar{T}(t)$, denoted $(z_1, y_1), \ldots, (z_k, y_k)$ in the dataset in saturation, we seek a non-increasing function $h : \mathbb{R} \to \mathbb{R}$ that estimates the mean $\bar{T} = g(\bar{Z}(t))$. The only constraint we impose is that the departure throughput is a non-increasing function of the number or runway crossings:

$$\min \sum_{i=1}^{k} (\hat{y}_i - y_i)^2 \tag{D.1}$$

subject to:

$$\hat{y}_i = g(z_i), \;\; i = 1, \ldots, k \tag{D.2}$$

$$g(i+1) \leq g(i), \;\; i = 0, \ldots (l-1), \text{where } l = \max(\bar{Z}(t)) \tag{D.3}$$

We do not impose more constraints because we are interested in observing the marginal cost of each crossing. From Figure D-6a we observe that the departure throughput given zero crossings is 4.5 departures/5 min, which implies a 66 sec inter-departure time. This is the same duration as the inter-departure time between flights without a runway crossing, which was estimated to be 66 sec in Section D.2.4 . For three crossings, the departure throughput decreases to 3.7 departures/5 min. This implies an 81 sec average inter-departure time. In Section D.2.4, it was found that the average inter-departure time given a runway crossing is 114 sec. The difference between these inter-departure times and the changing slope of the curve of Figure D-6a suggest that controllers tend to combine runway crossings. They are not likely to authorize a crossing after each departure, but instead conduct two crossings together during an inter-departure time.

We also note that the departure throughput decreases nonlinearly with increasing arrival crossings. It decreases at a higher rate when the number of crossings increases from 0 to 2, compared to when the number of crossings increases from 2 to 4 (when it stays stable). This suggests that the first two crossings are not likely to be combined, but for higher number of crossings, some of them will be performed during the same inter-departure time, thereby reducing their marginal cost. If this is the case, the balanced capacity implied from this figure lies between the two extremes considered before (no crossings combined, and all crossings combined in a staggered manner). Indeed, the balanced operations capacity of the runway is approximately 3.5 departures and 3.5 crossings/5 min. That would imply a capacity of 42 departures and 42 crossings/15 min.

For a more direct comparison with the results of the regression analysis, we parametrize the results with the number of Heavy aircraft. Because of the data sparsity, we classify B757's as Heavy aircraft and estimate the departure throughput as a function of runway crossings and Heavy departures. We define $H_{Deps}(t)$ as the number of B757's as Heavy aircraft that take off during the 5-minute interval $[t, t+5)$ min.

Given k triplets of measurements $\bar{Z}(t)$, $H_{Deps}(t)$ and $\bar{T}(t)$, denoted by $(z_1, w_1, y_1), \ldots, (z_k, w_k, y_k)$, we seek a function $g_h : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $\bar{T} = g_h(\bar{Z}(t), H_{Deps}(t))$. Thus, function g_h is a piecewise linear function of $\bar{Z}(t)$ and $H_{Deps}(t)$. The constraints are that the departure throughput is a non-increasing function of the number or runway crossings, and a non-increasing function of the number or Heavy aircraft in the fleet mix. The fitted function g_h is plotted in Figure

D-6b.

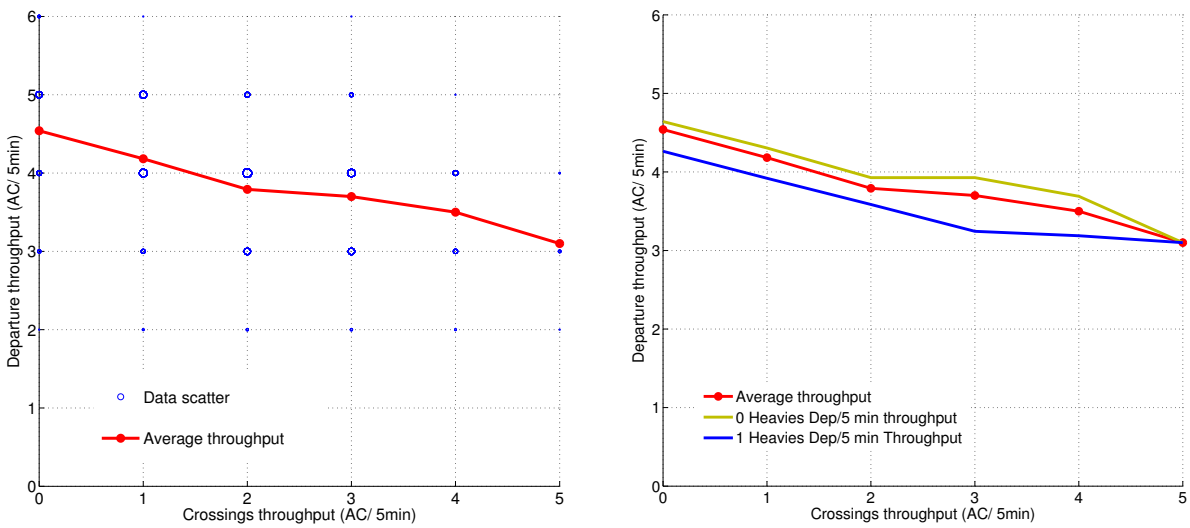
$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{D.4}$$

subject to:

$$\hat{y}_i = g_h(z_i, w_i), \;\; i = 1, \ldots, k \tag{D.5}$$

$$g_h(i+1, j) \leq g_h(i, j), \;\; i = 0, \ldots (l-1), \forall j, \text{where } l = \max(\bar{Z}(t)) \tag{D.6}$$

$$g_h(i, j+1) \leq g_h(i, j), \;\; j = 0, \ldots (n-1), \forall i, \text{where } n = \max(H_{Deps}(t)) \tag{D.7}$$



(a) Operational throughput envelope for Runway 17R.

(b) Operational throughput envelope for Runway 17R parametrized by the number of departing Heavies.

Figure D-6: Operational throughput envelopes for Runway 17R.

From Figure D-6b, we observe that the estimated curves are in agreement with the results of Section D.2.4. The departure throughput given one crossing is approximately equal to the departure throughput given one Heavy departure. Secondly, as the number of crossings increases, the departure throughput of 0 Heavy and 1 Heavy departures converges to the same number (3.1 departures/ 5 min). As the number of crossings increases, the controllers perform some crossings after a Heavy departure. In this way, the cost of one Heavy departure is on average 20 sec. By contrast, at lower numbers of crossings, there is a high probability that the crossing will not be performed after the Heavy departure. Thus, the departure throughputs given no Heavy departures and one Heavy departure decrease at the same rate for a low number of crossings (0-2).

The results of Figure D-6b confirm the results of Section 2.6.3, where we modeled the dependence of the departure throughput of BOS configuration 22L, 27 | 22L, 22R and concluded that as the number or arrival throughput (and thus arrival crossings) increases, the cost of Heavy departures becomes negligible. Here, using a much more detailed and accurate dataset, and relaxing all constraints in the estimation problem regarding the interactions between Heavy departures and crossings, we arrived at the same conclusion.

D.2.6 Estimation of DFW capacity

The results from the previous section can be also used for deriving approximate estimates of the capacity of the whole airport under the most frequent runway configuration 13R, 17C, 17L, 18R | 17R, 18L, 13L (props) . We initially focus on operations on the east side, namely, on arrival runways 17C, and 17L and departure runway 17R. We have not considered operations on the arrival runways at all, but the crossings observed at Runway 17R clearly provide a lower bound for the capacity of these runways. If 5 crossings can be performed through Runway 17R in 5 minutes, this implies that there is a demand for at least 5 crossings at Runway 17R, and that at least 5 aircraft have landed on Runways 17C and 17L.

For zero arrivals, the average departure throughput of Runway 17R is 4.5 AC/5 min, or 54 AC/hr. For balanced operations, the average throughput is 42 departures/hr and 42 arrivals/hr. Finally, for arrival priority, 5.0 crossings and 3.1 departures can be performed in 5 min, implying a capacity of 37 departures/hr and 60 arrivals/hr.

Given that operations on Runways 13R, 18R and 18L are symmetric to those on Runways 17C, 17L and 17R, we simply multiply the previous capacity estimates by 2 and obtain capacity estimates for configuration 13R, 17C, 17L, 18R | 17R, 18L. Departure priority capacity is 108 departures/hr, balanced capacity 84 departures/hr and 84 arrivals/ hr, and arrival priority capacity is 74 departures/hr and 120 arrivals/hr.

Finally, the departure capacity is increased by the prop departures on Runway 13L. In 2010, props comprised 5% of the fleet mix of the departing aircraft. Assuming that all of them are assigned to Runway 13L, we can approximately estimate the departure capacity of 13L. For every 19 jet departures of Runways 17R, 18L, there is one prop departure of Runway 13L. This hypothesis can be validated by comparing the saturation plot of Figure D-1 with the jet saturation plot of Figure D-7. The saturation throughput decreases from 20 to 19 when considering jets only. We have the following capacity estimates for runway configuration 13R, 17C, 17L, 18R | 17R, 18L, 13L

(props): Departure priority capacity is 116 departures/hr, balanced capacity is 88 departures/hr and 84 arrivals/hr, and arrival priority capacity is 78 departures/hr and 120 arrivals/hr. The AAR of 13R, 17C, 17L, 18R | 17R, 18L, 13L (props) is 126 operations/hr, so our estimate for the arrival priority arrival capacity is close to the declared airport acceptance rate.

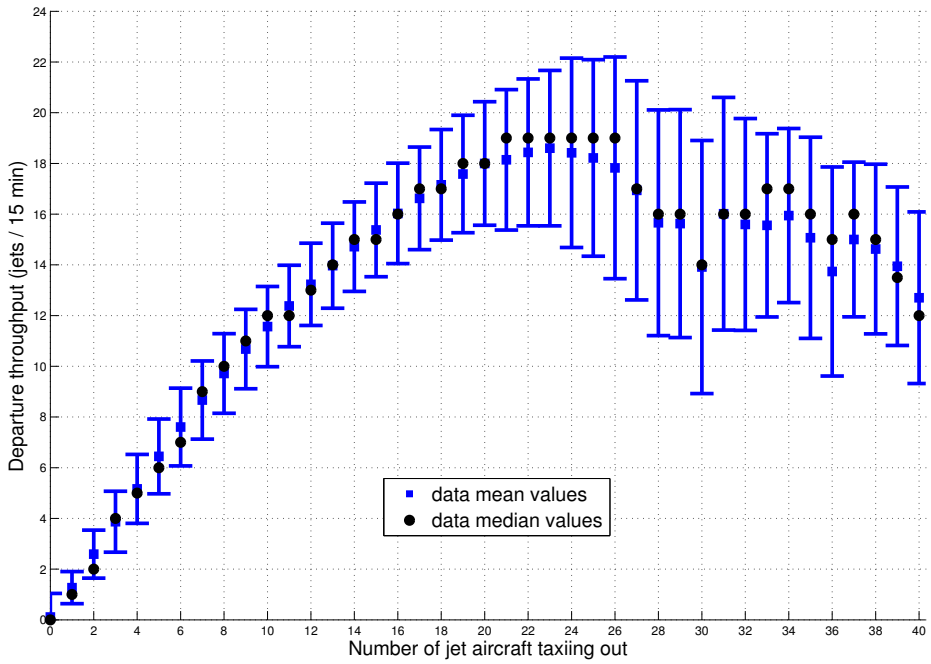


Figure D-7: Jet aircraft departure throughput as function of jets departure demand in 2010

In addition, 78 AC/hr is a lower bound for the available departure capacity of Runways 17R, 18L, and 13L. From Figure D-1, we note that the average departure throughput fluctuates between 13.5 and 19.5 AC/15 minutes, or 54 and 78 AC/hr. Thus, the average departure throughput of Runways 17R, 18L, and 13L takes values less than or equal to the lower bound of the available departure capacity. This appears to confirm our initial hypothesis, that the reason for the irregular shape of the saturation throughput curve is an inefficient utilization of the three departure runways.

D.3 Departure capacity of Runway 27L of PHL

D.3.1 Comparison of ASPM and ASDE-X data

In most of the applications considered in Chapters 2 and 3, arrivals cross the departure Runway, for example Runway 22L at BOS, and Runway 22L and 4R at EWR. Similarly, in Section D.2, we concluded that runway crossings are a major driver of the performance of Runway 17R. In all these

cases, it was observed that the impact of Heavy departures diminishes when there is a significant number of arrival crossings.

In this section, we study the departure capacity of Runway 27L, of the 26, 27R, 35 | 27L, 35 runway configuration, which is the most frequently one at PHL, and was used for 74% of the time in 2011. In Figure D-8, we show the saturation curve for this runway configuration using both ASPM data from year 2011 and ASDE-X data from the months June-August 2011. We notice that the curves are shifted from each other. This results from late ASDE-X transponder capture, as discussed in Sections 2.5 and 5.2.1. We also notice that the throughput in saturation is higher when measured with ASPM data. A probable explanation for this is that the ASDE-X data is only from the three summer months, and the high convective weather activity in the Northeastern US during this period resulted in reduced capacity [90].

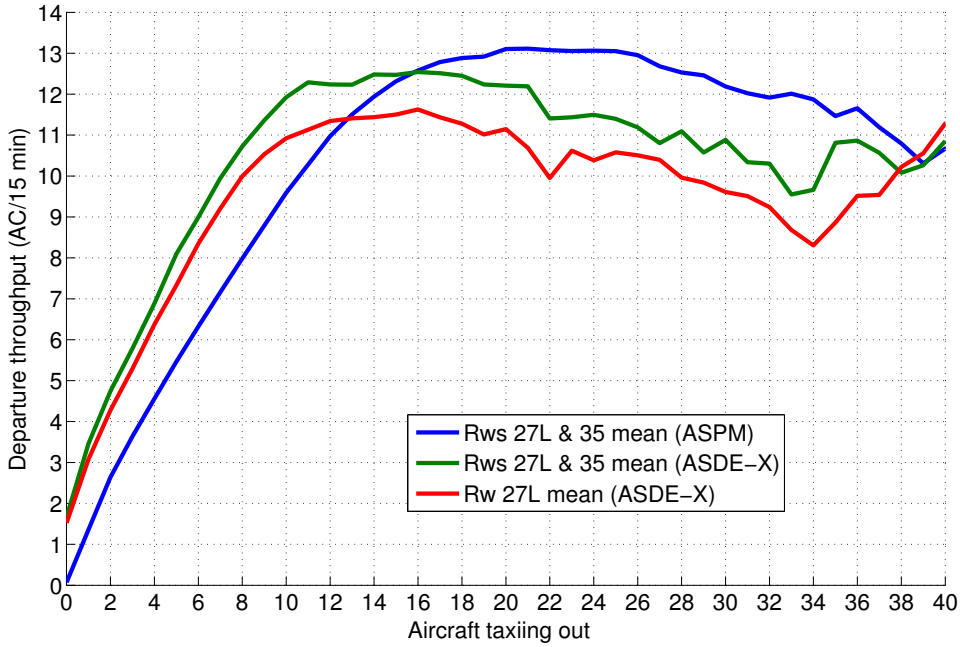


Figure D-8: Departure throughput as function of aircraft taxiing out at PHL.

D.3.2 Use of RAPT for estimating the operational throughput envelope of PHL

Using ASDE-X data, we isolate the saturation curve for the major departure runway 27L only, which we also show in Figure D-8. By inspection, we notice that in saturation, there is one departure from Runway 35 for every 11 to 12 departures from Runway 27L. We also observe that none of the curves fluctuates around the capacity in saturation, but instead, they all exhibit a clear decreasing trend. A possible explanation for this trend is that excessive congestion creates bottlenecks and

gridlocks. As a result, the performance of the airport decreases. Another reason could be that the throughput decreases as a function of other (hidden) variables, such as, route availability, or traffic management initiatives.

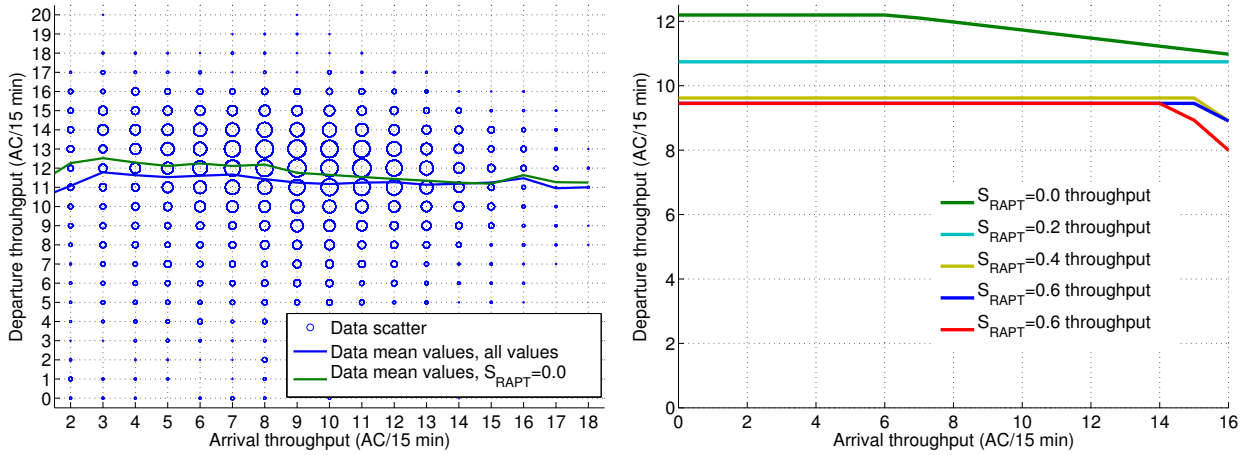
To derive the saturation area, we construct the regression tree that represents the departure throughput of Runway 27L as a function of its demand (the number of aircraft taxiing out that takeoff from Runway 27L), the arrival throughput, and the S_{RAPT} value, as calculated for LGA in Section 3.3.1. We do not have the RAPT values for PHL, but hypothesize that the route availability of LGA correlates well with that of PHL. The the area for which the departure throughput does not change with departure demand is estimated as $12 \leq N \leq 21$. Values of $N > 21$ imply lower throughput than $12 \leq N \leq 21$, all else being equal.

Figure D-9a shows the observed mean values of the departure throughput at each value of the arrival throughput for all S_{RAPT} values, and for $S_{RAPT} = 0.0$. We clearly see that the mean departure throughput given $S_{RAPT} = 0.0$ takes higher values (11.8 versus 11.4 AC/15 min). In addition, it has lower standard deviation (2.7 versus 2.9 AC/15 min), similar to the LGA observations (Section 3.3.1). We also observe that the departure throughput of Runway 27L is insensitive to the arrival throughput, as hypothesized. The major arrival runway 27R is between the terminals and Runway 27L, and the secondary arrival runways, 26 and 35, do not interfere with the departures either.

We apply the methodology described in Section 3.3.1 to estimate the departure throughput as a function of arrival throughput and route availability. From Figure D-9b, we note the decreasing trend of throughput with route availability. We also note that the throughput given $S_{RAPT} = 0.0$ is the only one that exhibits some (small) tradeoff with arrival throughput. It decreases by 1 AC/15 min as the arrival throughput increases.

D.3.3 Operational throughput envelope parametrized by Heavy aircraft departures

The conditions $12 \leq N \leq 21$, and $S_{RAPT} = 0.0$ guarantee persistent departure demand and high route availability. Thus, the departure throughput curve given $S_{RAPT} = 0.0$ in Figure D-9 can be viewed as the operational throughput envelope of Runway 27L. We measure the impact of Heavy aircraft departures by formulating the following estimation problem: Given k triplets of measurements $A(t)$, $H_{Deps}(t)$ and $T(t)$, denoted by $(u_1, v_1, y_1), \ldots, (u_k, v_k, y_k)$, at times when $12 \leq N \leq 21$, $S_{RAPT} = 0.0$, we seek a function $g_h : \mathbb{R}^2 \to \mathbb{R}$ that estimates the mean $T = g_h(A, H_{Deps})$.



(a) Mean departure throughput for all S_{RAPT} values and for $S_{RAPT} = 0.0$.

(b) Throughput envelopes parametrized by route availability.

Figure D-9: Departure throughput of Runway 27L of PHL as a function of arrival throughput and route availability.

As before, the constraints are imposed only between neighboring points:

$$\min \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{D.8}$$

subject to:

$$\hat{y}_i = g_h(u_i, v_i), \;\; i = 1, \ldots, k \tag{D.9}$$

$$g_h(i+1, j) \leq g_h(i, j), \;\; i = 0, \ldots (l-1), \forall j, \text{where } l = \max(A(t)) \tag{D.10}$$

$$g_h(i+1, j) - g_h(i, j) \leq g_h(i, j) - g_h(i-1, j), \;\; i = 1, \ldots (l-1), \forall j \tag{D.11}$$

$$g_h(i, j+1) \leq g_h(i, j), \;\; j = 0, \ldots (n-1), \forall i, \text{where } n = \max(H_{Deps}(t)) \tag{D.12}$$

Inequalities (D.10) and (D.11) are analogous to those in the case of the capacity envelope, i.e., for a given number of departing Heavy aircraft, the departure throughput is a monotonically non-increasing, concave function of the arrival throughput. Inequality (D.12) ensures that for a given value of arrival throughput, the departure throughput decreases as the number of Heavy departures increases. We do not impose any other constraints in this fitting problem to avoid making further operational assumptions. Given that arrivals do not interact with departures in this configuration, it is hypothesized that they do not interact with Heavy departures, either.

The estimated function is shown in Figure D-10. We notice that the operational throughput of

Runway 27L stays above 12AC/15 min for a large range of values of the arrival throughput. This can be contrasted with the operational throughput envelopes of EWR (shown in Figures 3-3a, 3-5), which stay at 11AC/15 min for the same range of arrival throughput. In addition, the impact of Heavy departures is clearly seen in Figure D-10. As Heavy aircraft departures increase from 0 to 4 AC/15 min, the departure throughput decreases by approximately 2AC/15 min.

In Chapter 1, as well as in Section 5.8.5, it was observed that PHL faces an acute congestion problem and that Runway 27L is often under very high pressure. Given that we filter out the conditions not conducive to high departure throughput, such as high congestion states ($N > 22$), or low route availability, one would expect Runway 27L to be more efficient for zero Heavy departures. Because Runway 27L is not used for runway crossings and offers dispersal headings, its departure throughput for zero Heavies was expected to be closer to 15 AC/15 min. The congestion problem at PHL, in combination with potential opportunities to increase departure throughput, suggest a very exciting direction for future research.

Finally, in Appendix C, we presented the departure throughput estimation results conducted with ASPM data for runway configuration 26, 27R, 35 | 27L, 35, during year 2011. The results here are very similar, but the departure throughput estimates at all states were higher in Appendix C as they included operations of Runway 35 as well.

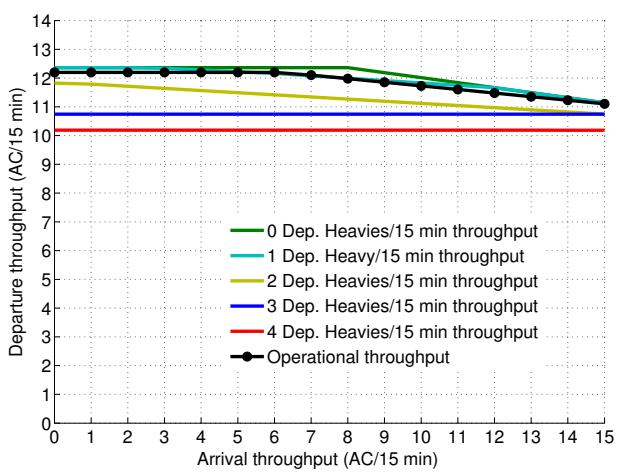


Figure D-10: Operational throughput envelope of Runway 27L of PHL parametrized by the number of Heavy departures.

Appendix E

Unimpeded Taxi-out Time Estimation Results

Table E.1: Unimpeded taxi-out time estimates for runway configuration 22L | 22R of EWR

Airline	Number of data points	Estimated mean	Estimated standard deviation
COA	19827	12.65	3.58
BTA	11141	12.33	3.26
CJC	3933	11.72	4.33
UAL	3333	12.26	3.02
UCA	3029	12.10	4.10
TCF	1982	11.66	3.58
DAL	1909	12.63	3.22
ASQ	2110	12.40	3.10
JBU	1723	15.35	4.02
SWA	1486	12.65	2.54
AAL	1270	12.65	3.28
USA	1231	13.32	4.07
ACA	986	12.40	3.75
PDT	563	8.63	3.39
EGF	528	11.38	2.34
DLH	496	12.55	4.17
other	8546	15.11	3.75

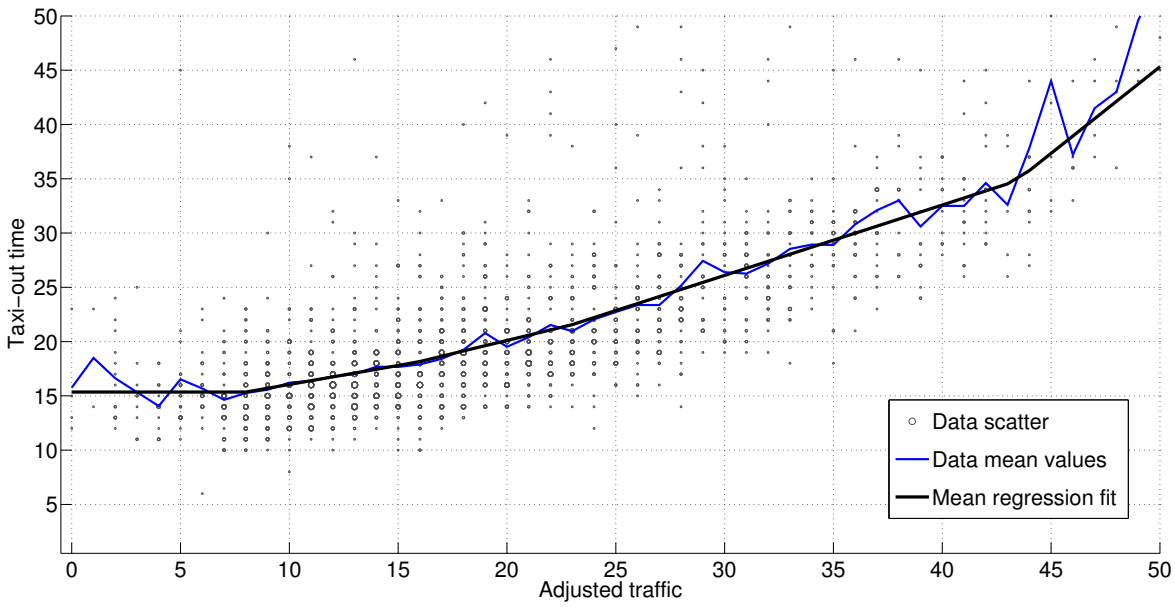


Figure E-1: Fitted function of the taxi-out times of the flights for JetBlue in configuration 22L | 22R of EWR

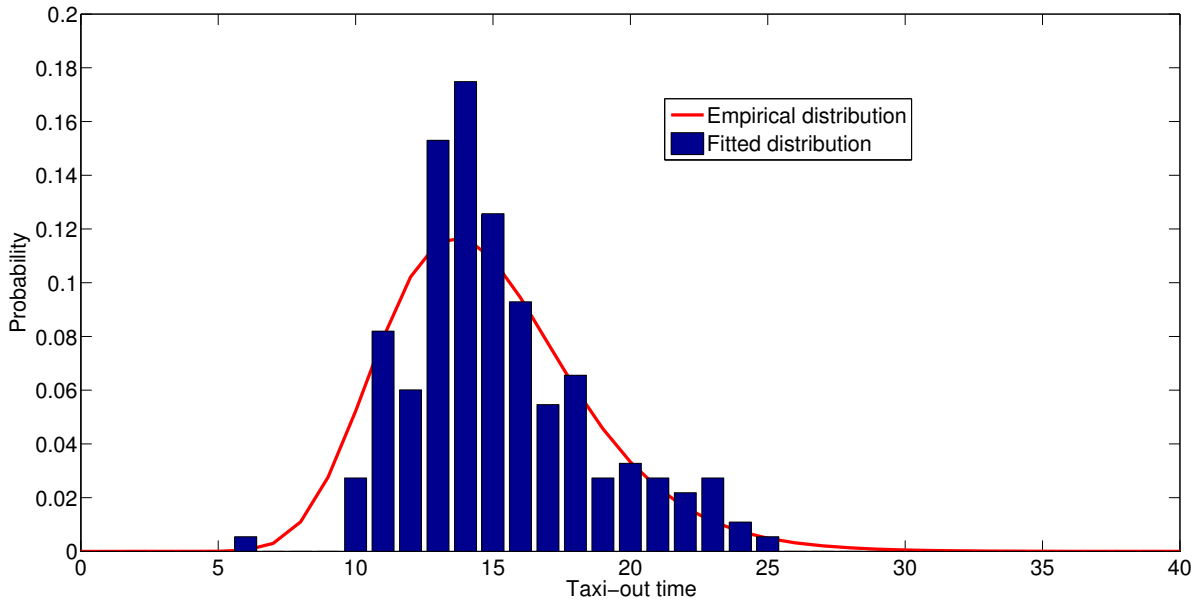


Figure E-2: Empirical and fitted distribution of the unimpeded taxi-out times for JetBlue in configuration 22L | 22R at EWR

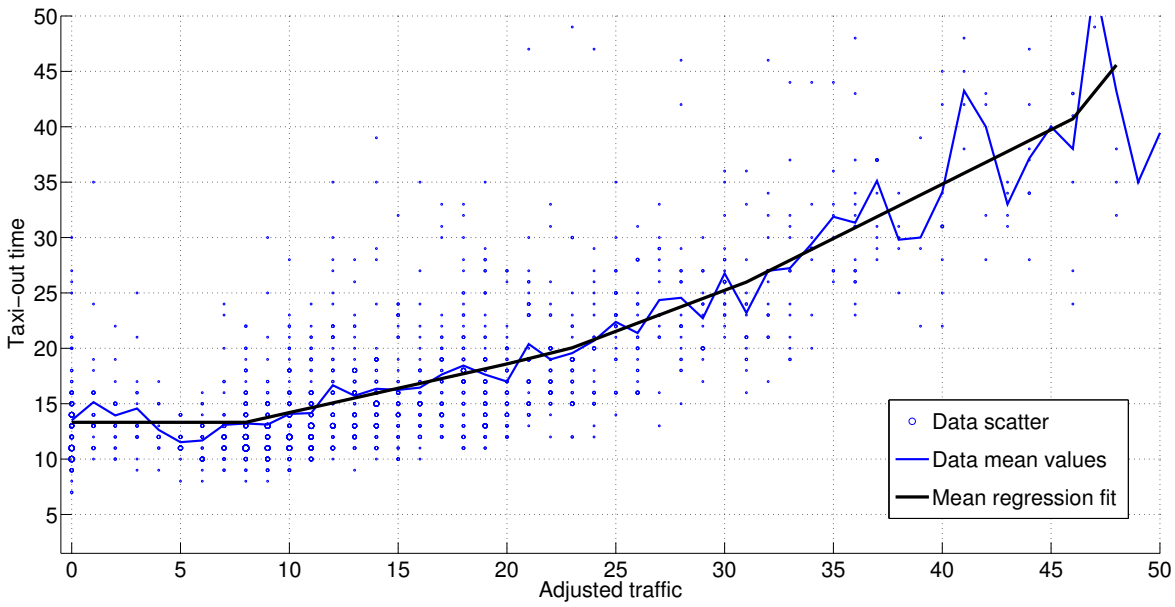


Figure E-3: Fitted function of the taxi-out times of the flights for US Airways in configuration 22L | 22R of EWR

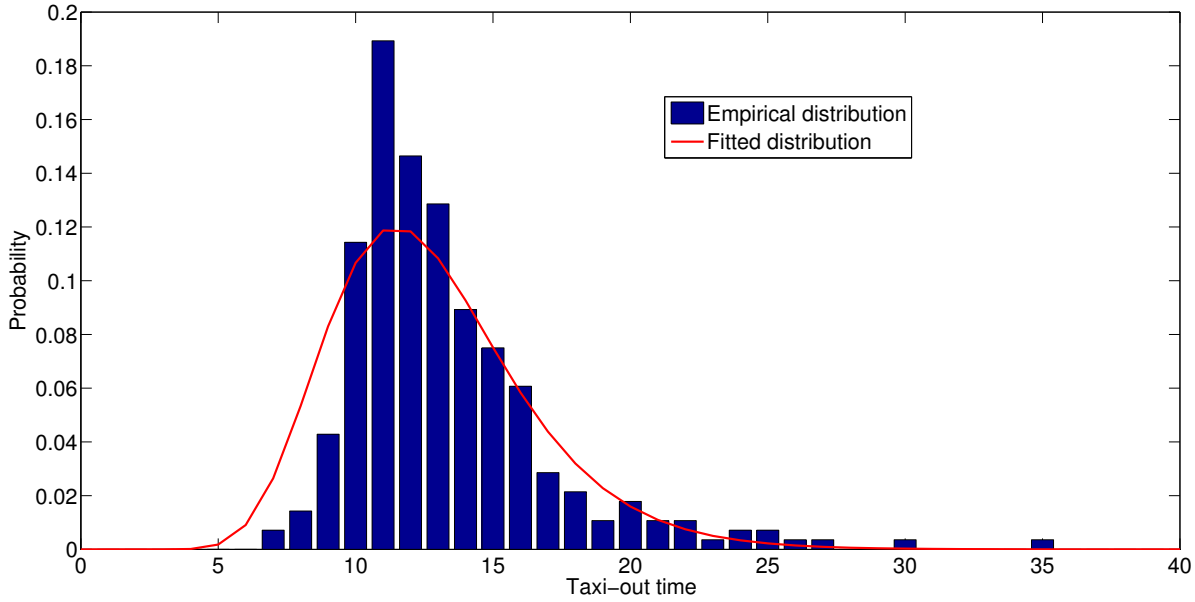


Figure E-4: Empirical and fitted distribution of the unimpeded taxi-out times for US Airways in configuration 22L | 22R at EWR

Appendix F

Model Predictions for EWR Runway Configuration 4R | 4L

F.1 Model development

Analogously to runway configuration 22R | 22L we develop the model using 2011 data from ASPM.

Table F.1: Aggregate taxi time predictions for EWR runway configuration 4R | 4L in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time	Mod. # of flights	Mod. mean taxi time
all	37132	22.73	37124	22.23
$(N \leq 8)$	13411	16.42	13113	16.16
$(9 < N \leq 14)$	11521	20.46	12246	20.33
$(N \geq 15)$	12200	31.80	11764	30.95

Table F.2: Prediction statistics for the congestion state and the throughput for EWR runway configuration 4R | 4L in year 2011.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.31	1.67	3.04	-1.15	2.75	4.38
Throughput (AC/15 min)	-0.01	1.10	1.55	-0.22	1.25	1.69

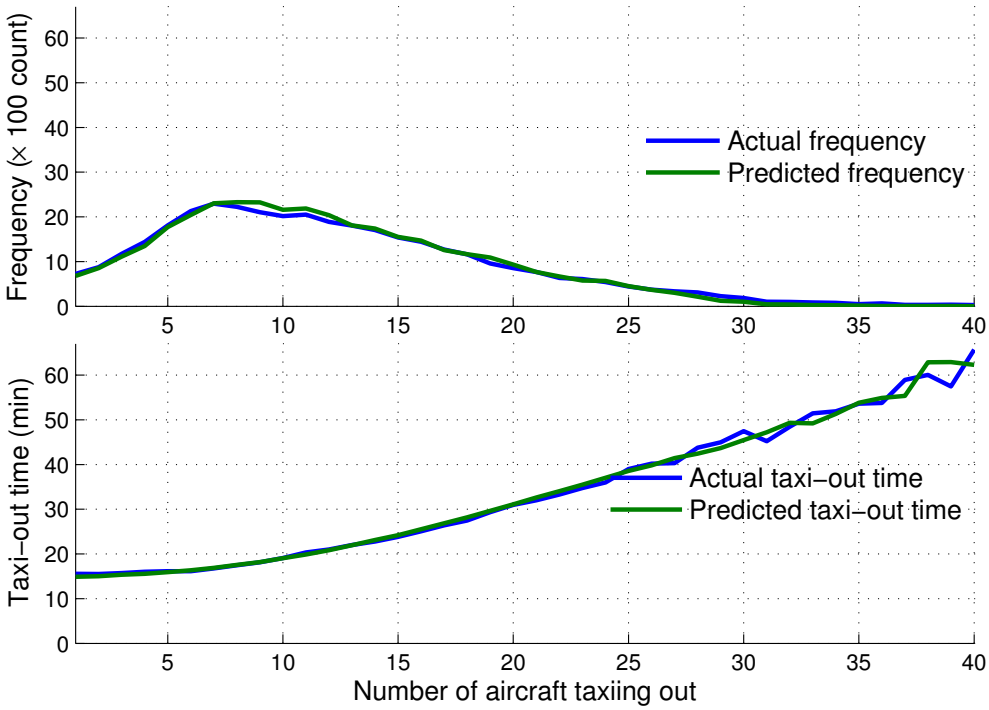


Figure F-1: Actual and modeled frequency of all states N (top); Actual and modeled dependence of the average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 4R | 4L in year 2011.

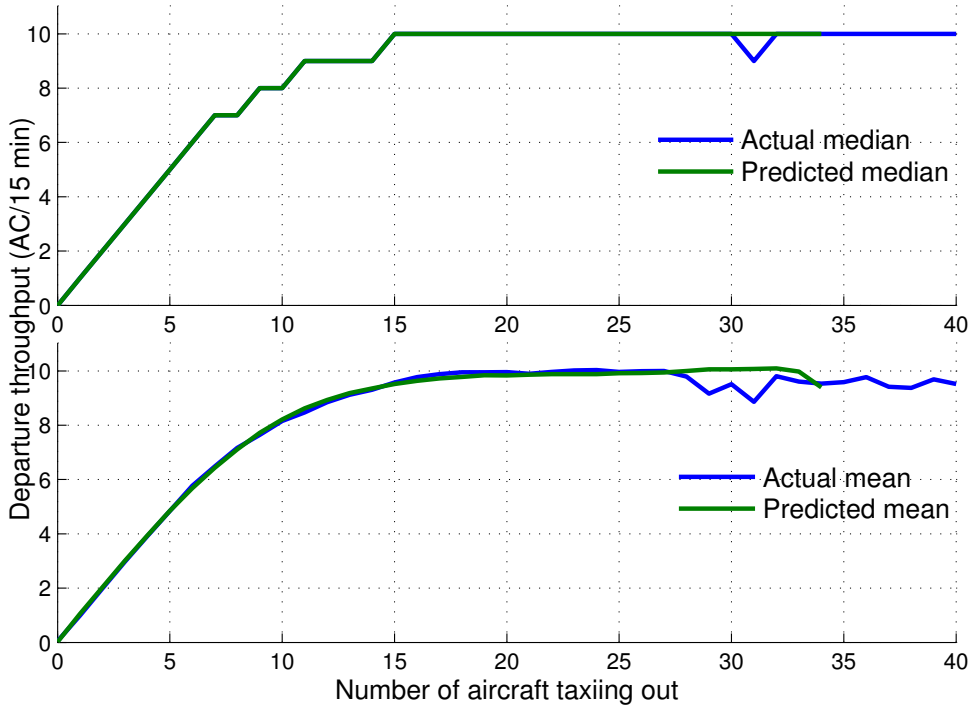


Figure F-2: Actual and modeled throughput of all states N for EWR runway configuration 4R | 4L in year 2011: Mean (top); Median (bottom).

F.2 Predictions for year 2010

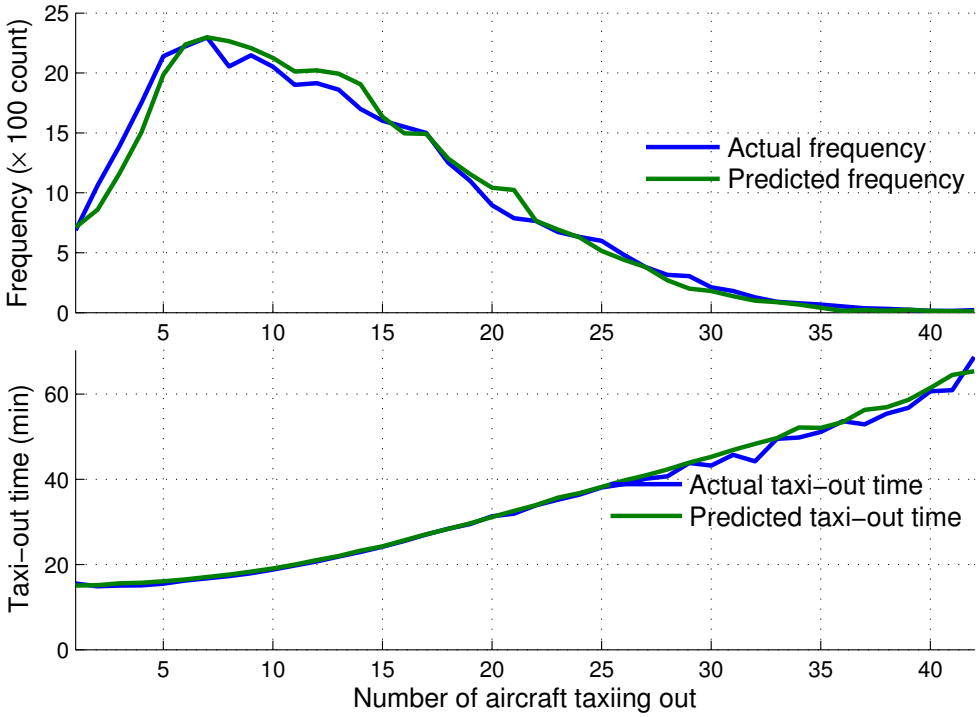


Figure F-3: Actual and modeled frequency of all states N (top); Actual and modeled dependence of the average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 4R | 4L in year 2010.

Table F.3: Aggregate taxi time predictions for EWR runway configuration 4R | 4L in year 2010.

Congestion level	Actual # of flights	Actual mean taxi time	Mod. # of flights	Mod. mean taxi time
all	39,785	22.86	29723	23.03
$(N \leq 8)$	14,285	15.94	13695	16.30
$(9 < N \leq 14)$	11,599	20.24	12255	20.55
$(N \geq 15)$	12200	32.17	13773	31.92

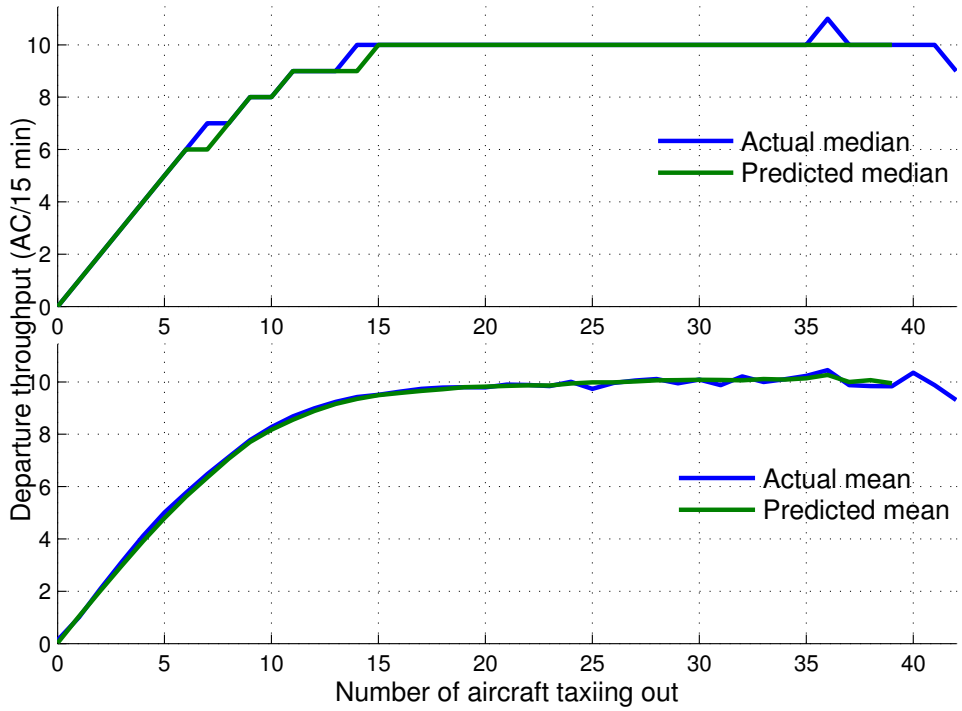


Figure F-4: Actual and modeled throughput of all states N: Mean (top); Median (bottom) for EWR runway configuration 4R | 4L in year 2010.

Table F.4: Prediction statistics for the congestion state and the throughput for EWR runway configuration 4R | 4L in year 2010.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.06	1.67	2.84	-0.70	2.63	3.97
Throughput (AC/15 min)	-0.03	1.10	1.55	-0.24	1.23	1.67

F.3 Predictions for year 2007

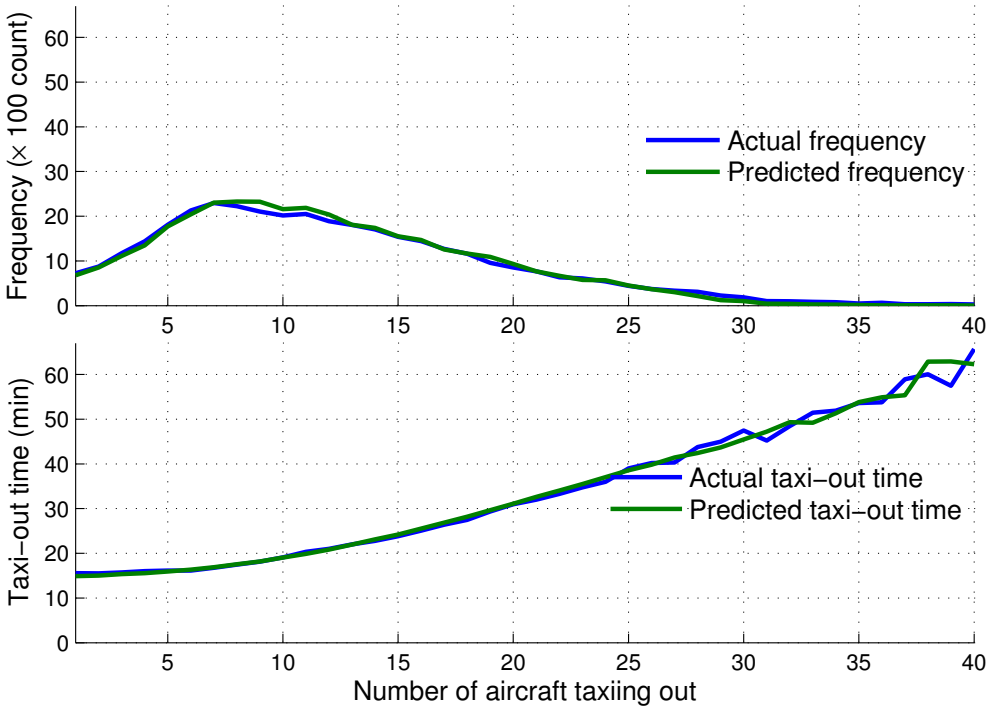


Figure F-5: Actual and modeled frequency of all states N (top); Actual and modeled dependence of the average taxi-out time as a function of the state N at the time of pushback (bottom) for EWR runway configuration 4R | 4L in year 2007.

Table F.5: Aggregate taxi time predictions for EWR runway configuration 4R | 4L in year 2007.

Congestion level	Actual # of flights	Actual mean taxi time	Mod. # of flights	Mod. mean taxi time
all	34,378	29.55	34400	29.43
$(N \leq 8)$	8,418	17,44	8751	16.57
$(9 < N \leq 14)$	7,861	21.60	8126	20.54
$(N \geq 15)$	18,099	38,61	17523	39.97

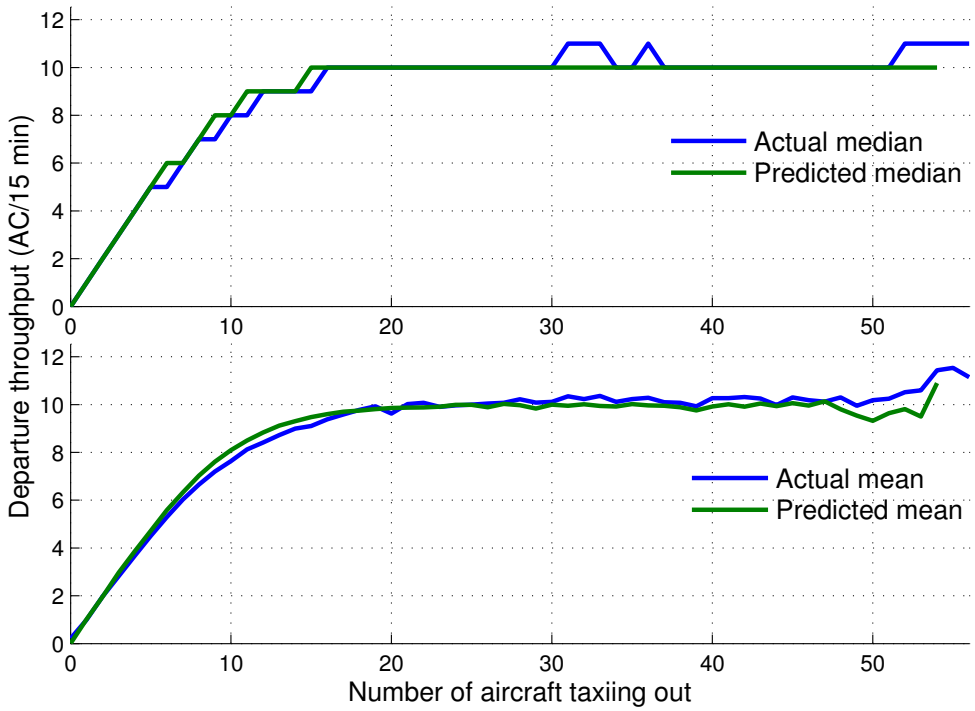


Figure F-6: Actual and modeled throughput of all states N for EWR runway configuration 4R | 4L in year 2011: Mean (top); Median (bottom).

Table F.6: Prediction statistics for the congestion state and the throughput for EWR runway configuration 4R | 4L in year 2007.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.25	2.58	4.17	-0.75	2.58	4.17
Throughput (AC/15 min)	-0.02	1.32	1.85	-0.23	1.42	1.94

F.4 Individual flights taxi-out times predictions

In addition to this, rather aggregate, comparison it is interesting to see how the model predicts individual taxi times, to compare the predicted taxi-out time for the flights out of EWR segment (VMC; 4R | 4L) in 2011, 2010 and 2007 with their recorded ones. Figure F-7 shows the cumulative distribution or the prediction error $E(i)$ defined as:

$$E = \tau(i)^{sim} - \tau(i)^{obs} \tag{F.1}$$

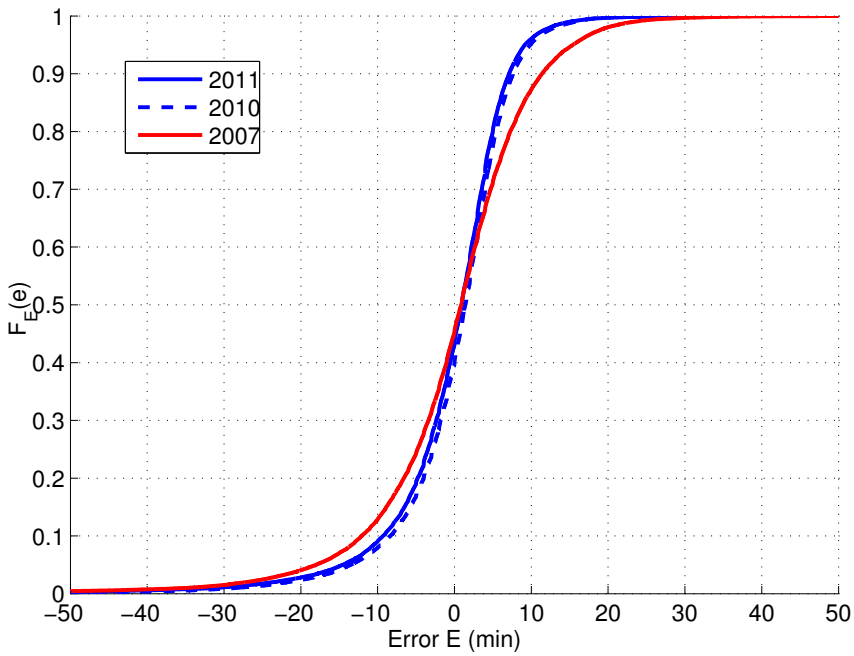


Figure F-7: Individual flights taxi time prediction error for EWR runway configuration 4R | 4L.

Table F.7: Prediction statistics for individual taxi-out times for EWR runway configuration 4R | 4L.

Year	ME	MAE	RMSE
2011	-0.50	5.58	8.67
2010	0.16	5.48	8.17
2007	-0.12	7.63	11.27

Appendix G

Model Predictions for PHL Runway Configuration 26, 27R, 35 | 27L, 35

In this appendix, we describe the departure process model for the most frequently used runway configuration at PHL, 26, 27R, 35 | 27L, 35. The model is applied in Section 5.8.5 for evaluating the performance of different congestion control strategies at PHL. For developing the model, we apply the methodology described in Chapter 4 using ASPM data from 2011, supplemented with data from Flightstats for obtaining terminal and gate information and gate-out information for the non-OOOI flights.

On the methodological front, we demonstrate how the runway service time distributions can be parametrized by the aircraft type of the departing aircraft. In the other model development examples considered in the thesis (EWR in Chapter 4 and CLT in Appendix H), the service time distributions depend on exogenous variables, like the route availability and the arrival throughput.

G.1 Saturation plot

For estimating the departure process characteristics, we follow the approach described in Section 4.4.3. We first depict the saturation plot for runway configuration (VMC; 26, 27R, 35 | 27L, 35), which is shown in Figure G-1. We note that the average throughput reaches its maximum value, 13 AC/15 min, for 20 aircraft taxiing out, and it starts decreasing at congestion states higher than 25 AC. The throughput decreases to 12 AC/15 min as the number of aircraft on the ground increases from 25 to 35 and decreases even further for 35 or more aircraft on the ground. We also note that the departure throughput exhibits very high variability.

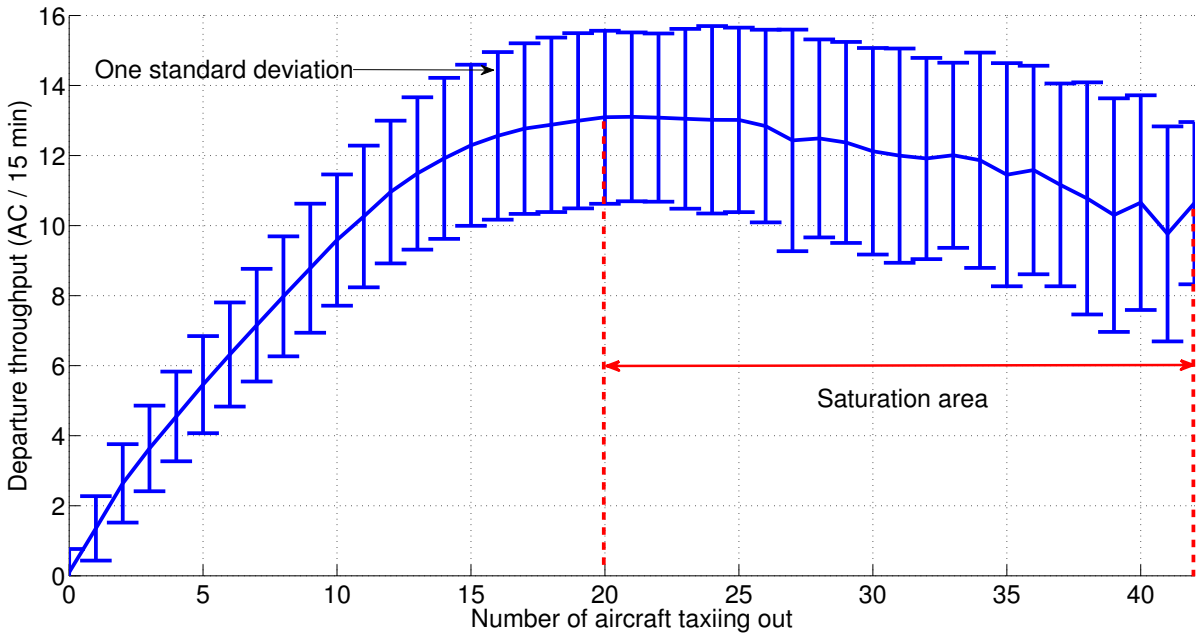


Figure G-1: PHL saturation plot for runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

For estimating the area, for which the departure throughput does not change significantly with the departure demand, we follow the approach outlined in Section 2.3. We use regression trees to estimate the departure throughput as a function of all potentially significant explanatory variables: Departure demand, route availability[1], arrival throughput, Heavies departures, and props departures. Route availability, arrival throughput, and Heavies departures were found to be significant variables in the analysis of ASDE-X data in Section D.3 of the appendix. Props departures are hypothesized to play a significant role, because props often use the secondary runway, Runway 35 [99]. A simplified version of the resulting regression tree is shown in Figure G-2. From the tree, we infer that the area that is not associated with increasing, or decreasing trend of the departure throughput is $20 \leq N \leq 27$.

G.2 Estimation of the departure capacity distributions

We use the saturation data-points to estimate the departure throughput as a function of the arrival throughput, the route availability and the fleet-mix. Consistently with the method presented in Section 4.4.3, we use regression trees. However, in the case of PHL, information about fleet mix shall be used for predicting the departure capacity. From Figure D-10, we note that the departure

[1]For route availability, we use RAPT data from LGA, as explained in Section D.3.

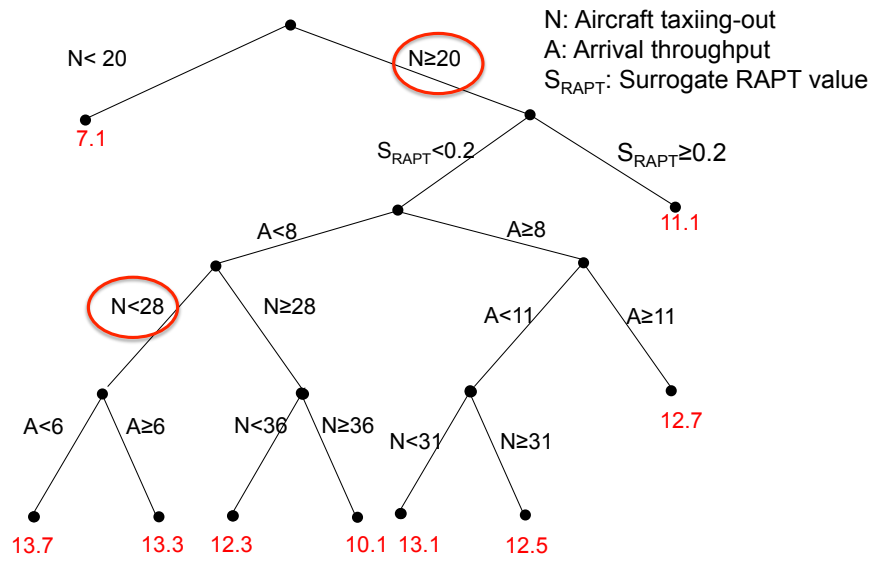
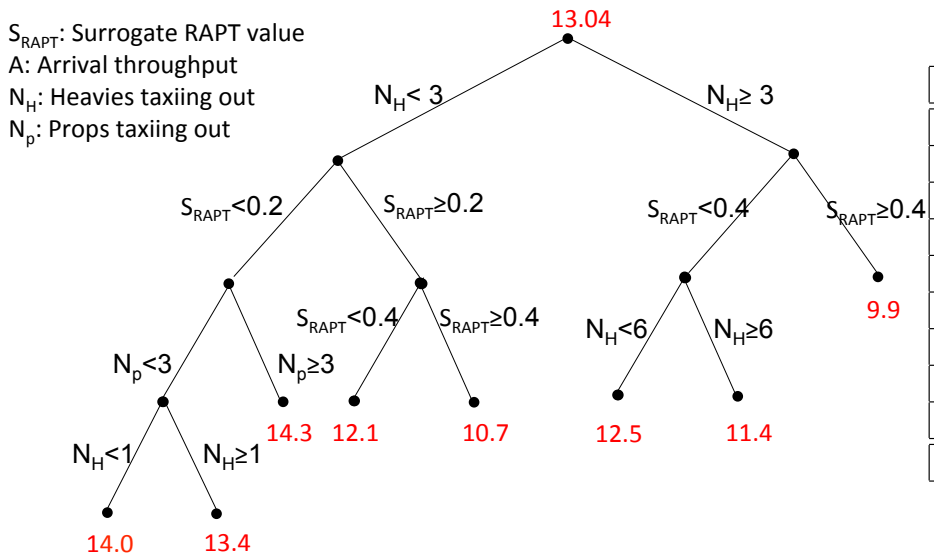


Figure G-2: Estimation of the area for which the departure throughput does not change with departure demand.

throughput decreases with the number of Heavy aircraft departing in a 15-minute window. We would like to use this information for dynamically predicting the departure throughput. This poses the difficulty that the number of aircraft departing, and, thus, the number of Heavy aircraft departing, is an output of the model. For this reason, we use information about the demand for departures from Heavies and props, that is the number of Heavy aircraft taxiing out (N_H) and the number of props taxiing out (N_p). The resulting regression tree, as well as the parameters of the fitted Erlang distributions are shown in Figure G-3. From the regression tree we note that the higher values of Heavy aircraft taxiing-out are associated with lower departure throughput. This is consistent with the capacity envelope of PHL, parametrized by the number of Heavy departures, shown in Figure D-10.

Conversely, higher values of prop aircraft taxiing-out are associated with higher departure throughput. This was expected, because props tend to use the secondary departure runway, and thus increase the overall departure throughput. Furthermore, decreasing route availability is associated with decreasing departure throughput for all fleet mix conditions. We finally note that the arrival throughput turns out to be a non-significant variable for predicting the departure throughput. This was expected, because the main arrival runway, 27R, is between the main terminal area and the main departure runway, 27L, and thus arriving traffic does not interfere with operations at the main departure runway. The operational throughput envelope for this runway configuration,

shown in Figure C-3, also indicates the very small dependence of the departure throughput on the arrival throughput.



Forecast	μ	k
1	9.9	1
2	14.3	2
3	12.1	2
4	10.7	1
5	12.5	3
6	11.4	3
7	14.0	3
8	13.4	3
8	13.4	3
All	12.9	2

Figure G-3: Expected departure throughput and service time distributions parameters for $20 \leq N \leq 27$ conditioned on arrival demand, route blockage and fleet mix information.

G.3 Unimpeded taxi-out estimation

For estimating the unimpeded taxi-out times, we adapt the method developed in Section 4.3.1. US Airways (USA) uses gates from terminals A, B, and C. Similarly, Republic Airlines (RPA) uses gates from terminals B and C. All these gates are spread over a large area of the airport. For this reason, we estimate the unimpeded taxi-out times of the flights of US Airways and Republic Airlines by terminal. In addition, all international flights of US Airways depart from terminal A. For these flights, we estimate the unimpeded taxi-out time separately, because they tend to have different procedures [25, 110] and move at slower speed. The results can be seen in Table G.1. We note that the international flights of the US Airways (USA) have longer unimpeded taxi-out times from the rest of the US Airways flights of terminal A. The average unimpeded taxi-out time of British Airways (BAW) flights is very similar. Piedmont Airlines (PDT), the fleet of which comprises only turboprops, has very short unimpeded taxi-out time. We also observe, that airlines using terminals D and E (DAL, AAL, SWA, TRS, EGF, UAL and COA) which are the closest to the threshold of Runway 27L have shorter unimpeded taxi-out times than airlines that use terminal A, B, C and F (USA, AWI, CHQ, RPA and BAW).

Table G.1: Unimpeded taxi-out time estimates for PHL runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

Airline	Estimated mean (min)	Estimated standard deviation (min)
USA, terminal A, domestic	14.03	4.65
USA, terminal A, international	15.56	4.60
USA, terminal B	14.14	5.68
USA, terminal C	11.37	3.80
AWI	11.89	3.19
PDT	9.03	2.77
RPA, terminal B	12.25	3.90
RPA, terminal C	12.76	4.08
SWA	9.01	1.73
DAL	11.51	3.28
AAL	10.81	3.55
UAL	10.69	2.61
CHQ	12.15	3.58
TRS	10.25	2.09
EGF	10.07	2.72
COA	9.88	2.70
LOF	10.51	1.86
ACA	11.66	2.46
MES	10.18	2.32
UCA	13.22	4.47
ASQ	10.40	2.16
COM	13.67	4.06
BAW	14.92	2.40

G.4 Results

In this section, we discuss the prediction results for the most frequently used runway configuration of PHL, (VMC; 26, 27R, 35 | 27L, 35), in 2011. The unimpeded taxi-out times estimated parameters are listed in Table G.1 and the service time parameters are shown in Figure G-3. As explained in Section 4.3.2, α is calculated so that the predicted median taxi-out time equals the actual median taxi-out time (16 min) and equals 0.22 min/AC.

This model is only different from that described in Section 4.4.5 of Chapter 4 in its use of endogenous information for predicting the departure capacity in each 15-minute period. At the beginning of each 15-minute period, the model first predicts the number of props and Heavies on the ground and then chooses the appropriate service time distribution (Figure G-3). This is accomplished by estimating the expected number of Heavies and props taxiing out following the method outlined in Section 4.4.5. The takeoff time of each aircraft is assumed to be $C_l + \tilde{d}_l(j)$, that is, the sum of its arrival time at the departure queue and is effective queuing delay. Repeating this calculation for all Heavy and prop aircraft yields the expected takeoff schedule for these aircraft. From their expected takeoff times, we calculate the expected number of Heavies and props on the ground, which are then used to derive the service time distributions for each 15-minute interval.

Figure G-4 shows the frequency of the different congestion states observed in the operational data and predicted by the model. The model predicts the airport being as often as observed at the higher congestion states ($N > 15$). However, the model underpredicts the number of aircraft that push back in medium congestion ($5 < N \leq 15$) and correspondingly it overpredicts the number of aircraft in low congestion ($N \leq 5$). We hypothesize that this is an artifact of the runway utilization. As shown in Figure D-8, in saturation, there is one departure out of runway 35 for every 11 to 12 departures out of runway 27L. However, at lower congestion states the secondary runway may be used even less frequently.

Figure G-4 also shows the expected taxi-out time as a function of the number of aircraft taxiing-out at the time of pushback for both the actual and the modeled operations. Consistently with the traffic state predictions, we note that the taxi-out time is predicted very accurately for $N > 15$. The underprediction of congestion results in a slight underprediction of taxi-out times at congestion states lower than 15.

Table G.2 contains more detailed statistics about the number of aircraft and the taxi times in different congestion levels. In agreement with the plots of Figure G-4, the model predicts accurately

both the frequency of the different congestion states and the taxi-out times at higher congestion states ($N \geq 15$).

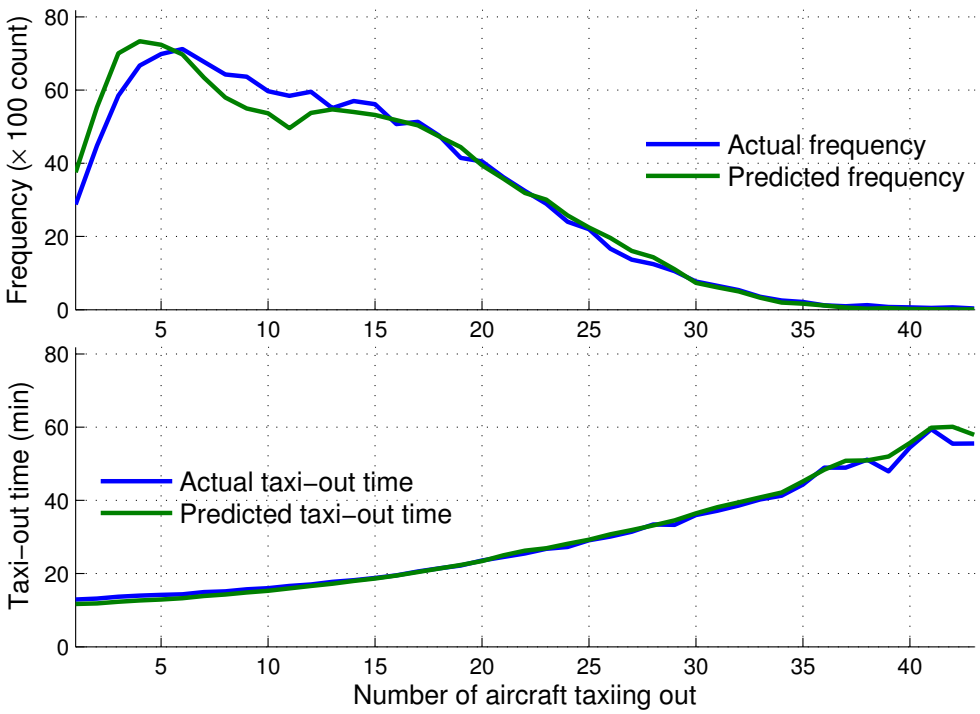


Figure G-4: Actual and modeled frequency of all states N (top); Actual and modeled dependence of the average taxi-out time as a function of the state N at the time of pushback (bottom) for PHL runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

Table G.2: Aggregate taxi time predictions for PHL runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

Congestion level	Actual # of flights	Actual mean taxi time (min)	Mod. # of flights	Mod. mean taxi time (min)
all	136286	18.98	136334	18.30
($N \leq 8$)	48874	14.18	52041	12.89
($9 < N \leq 14$)	35361	16.85	32064	16.343
($N \geq 15$)	**52051**	**24.96**	**52207**	**24.86**

In Figure G-5, we show the mean and median throughput as predicted by the model as a function of the congestion state N. The model predicts both the mean throughput and the median throughput very accurately in higher traffic conditions ($N \geq 15$). It is noteworthy that the model predicts the decrease of the mean and median departure throughput at very high congestion states ($N \geq 30$). This is because higher congestion states are associated with many Heavies, fewer props and low route availability, which imply lower departure capacity (Figure G-3). The predictions for

the 15-minute throughput at each minute are also listed in Table G.3. This case demonstrates the importance of dynamic service time distributions. If we used only one service time distribution, this would be derived from the departure throughput in saturation ($20 \leq N \leq 27$), which implies service time Erlang distributed with parameters (2×13.0, 2). The results from applying the stochastic and static model (S.M.) are also shown in Figure G-5 with the dashed black line. We note, that, as expected, the static service time distributions predict a steady departure throughput at 13 AC/15 min after the airport enters the saturation regime.

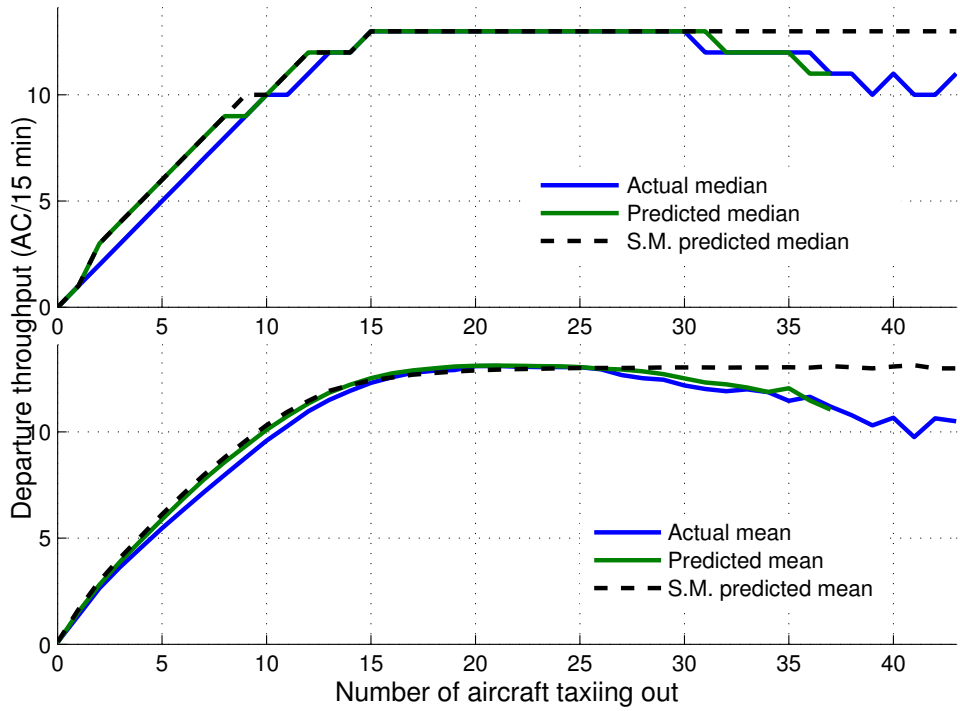


Figure G-5: Actual and modeled throughput of each state N: Mean (top); Median (bottom) for PHL runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

Table G.3: Prediction statistics for the congestion state and the throughput for PHL runway configuration 26, 27R, 35 | 27L, 35 in year 2011.

	$N(t) > 0$			$N(t) \geq 10$		
	ME	MAE	RMSE	ME	MAE	RMSE
State (AC)	-0.45	1.85	3.32	-0.79	2.99	4.79
Throughput (AC/15 min)	-0.01	1.48	2.07	-0.2	1.82	2.42

We can use the model to predict the evolution of the departure throughput and taxi-out times over a day at PHL. In the upper plot of Figure G-6, we show the average number of pushbacks

and the average number of takeoffs (or departures) that was recorded during each 15 minute time-window for all days in which this runway configuration was in use in 2011. We also show the average number of departures of this runway configuration of PHL, as predicted by the model. In the lower part of Figure G-6, we show the actual and predicted average taxi-out times for the flights that pushed back in each 15-minute time window. We observe that the model is representative of an average day at PHL.

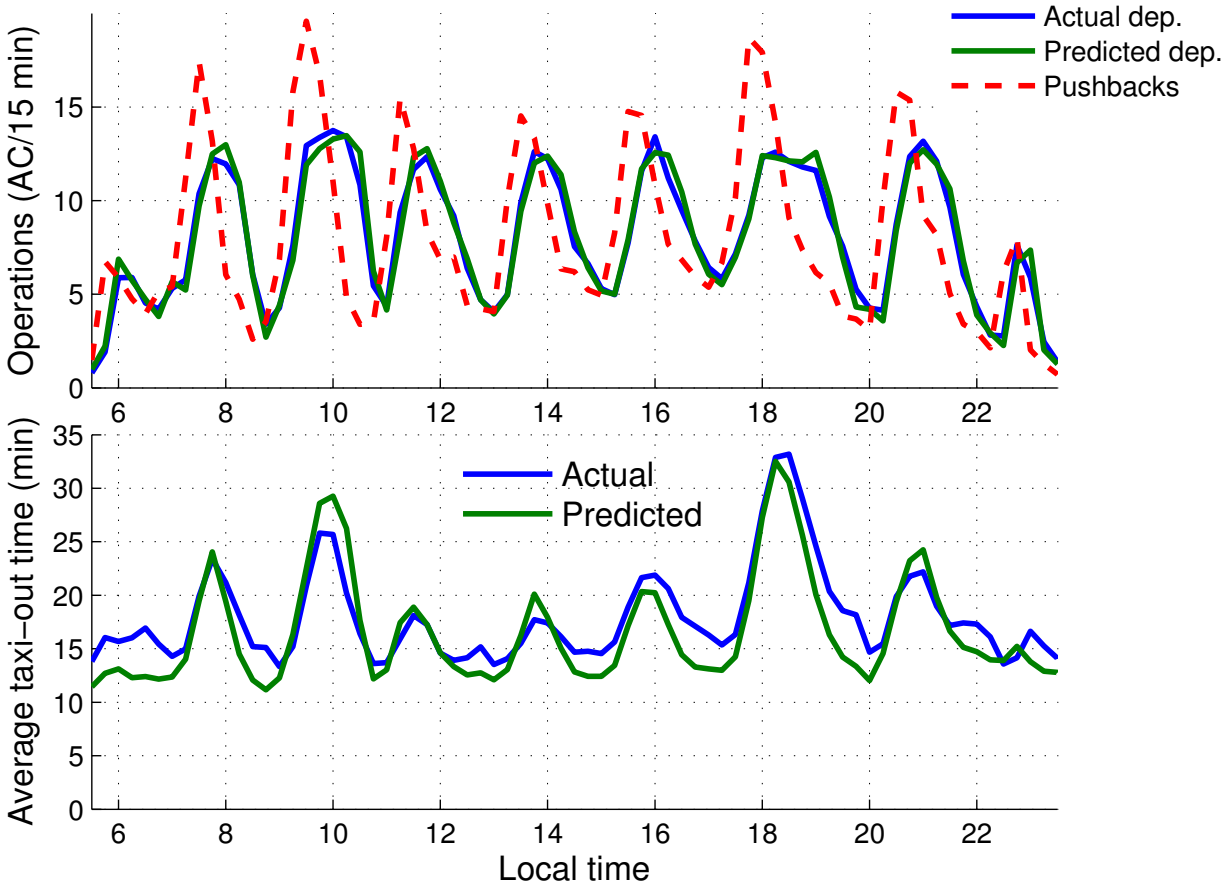


Figure G-6: Average number of pushbacks, average number of actual and predicted takeoffs at PHL in 2011 (top); Average actual and predicted taxi-out times (bottom).

Finally, in Figure G-7, we show the predictions for the departure throughput of Heavies and props averaged over all days that this runway configuration was in use in 2011. We notice that the departure throughput of both Heavy aircraft and props is predicted very accurately over the course of the day. Thus, we see that the model not only uses fleet mix information to predict the total departure throughput accurately, as shown in Figures G-5 and 5-17, but it also accurately predicts the throughput of the aircraft types which are expected to reduce (Heavy aircraft) or

increase (props) the departure throughput. This means that, on average, Heavy aircraft and props are predicted to stay on the ground for as long as they actually do and thus using information about their expected number on the ground does not bias the estimates. Concluding, we note by inspecting Figures 5-17 and G-7 that the low departure throughput and the high taxi-out times associated with the evening departure push (1730 to 1930 hours) are partly explained by the large number of Heavy aircraft and the small number of props pushing back in that time-period.

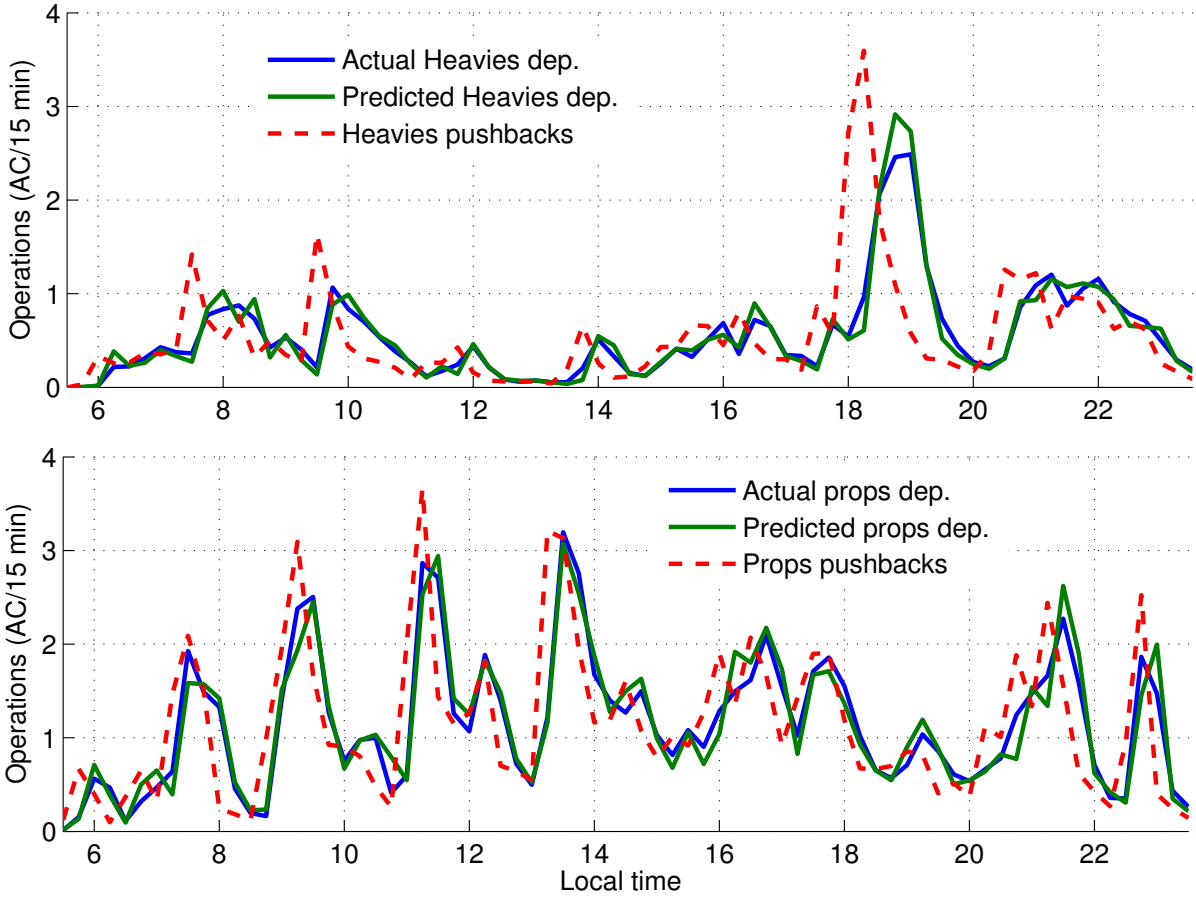


Figure G-7: Average number of pushbacks, average number of actual and predicted takeoffs of Heavies at PHL in 2011 (top); Average number of pushbacks, average number of actual and predicted takeoffs of props at PHL in 2011 (bottom).

Appendix H

Assessment of the Impacts of the New Runway at Charlotte International Airport

In this appendix, we present a case study with policy implications. First, we apply the operational throughput method presented in Chapter 2 to measure the capacity increase in Charlotte International Airport (CLT) resulting from the construction of runway 18R/36L in January 2010. Subsequently, we apply the analytical queuing model presented in Chapter 4 to estimate the taxi-out delay savings due to the added runway.

In addition to assessing of the operational gains from the deployment of the new runway, this appendix offers two significant methodological contributions. Firstly, we show that operational throughput envelopes can differ not only in counts, but also in shape from theoretical estimates, and secondly, we show that the queuing model can be successfully applied to multi-departure runway systems.

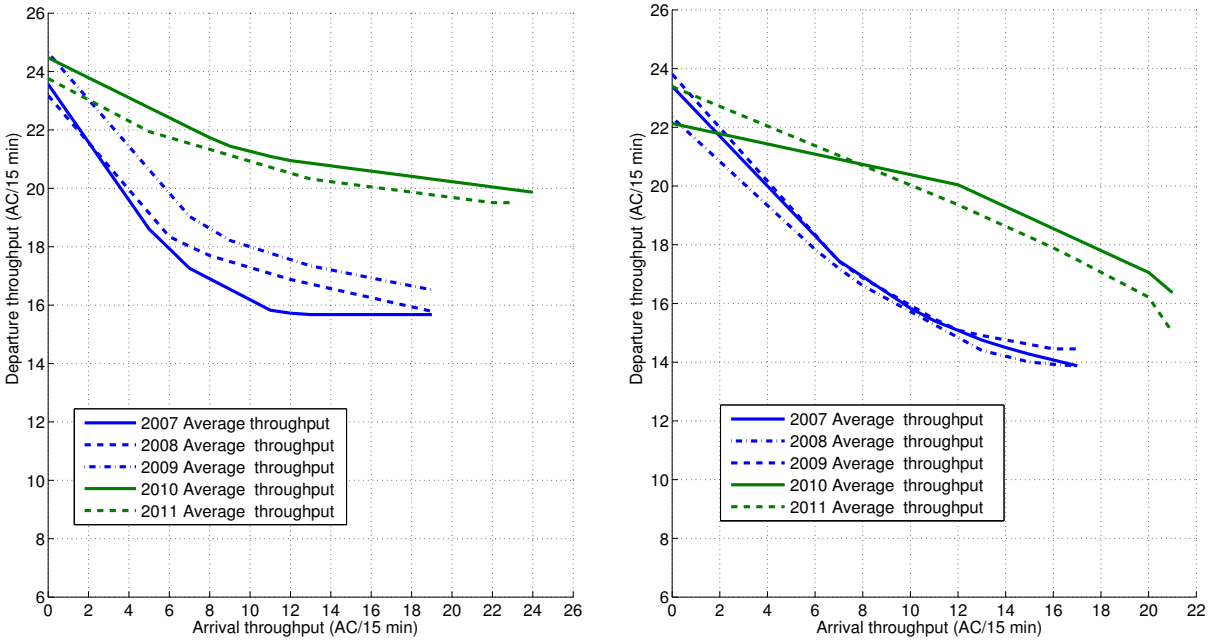
H.1 Operational throughput envelopes at CLT before and after the new runway

Runway 18R/36L was constructed in 2009-2010, and became operational in January 2010. The airport diagram can be seen in Figure A-6. The new runway added significant capacity to the two major runway configurations of the airport in use during VMC:

- South flow configuration: 18R, 23 | 18R, 18L became 18C, 18R, 23 | 18C, 18L.

- North flow configuration: 36L, 36R | 36L, 36R became 36C, 36L, 36R | 36C, 36R.

In Figure H-1, we present the operational throughput envelopes for the two major runway configurations for three years before and two years after the addition of the new runway. The capacity estimates are also summarized in Tables H.1 and H.2.



(a) Operational throughput envelope for the south flow configurations.

(b) Operational throughput envelope for the north flow configurations.

Figure H-1: Operational throughput envelope of the major CLT runway configurations.

In Figure H-1, we notice two families of curves for both runway configurations: Those from the years before the addition of the new runway (2007, 2008, and 2009) and those after (2010, and 2011). The departure capacity appears to have increased by 2 to 4 AC/15 min for a given arrival throughput and the arrival priority capacity by 4 AC/15 min. For the case of the north flow configuration, shown in Figure H-1b, we also note the stability of the estimated curves, that is, the estimated operational throughput envelopes are almost identical for the years 2007-2009 and 2010-2011. There is a higher variability in the estimated operational throughput envelopes for the south flow runway configuration, shown in Figure H-1a, in the years 2007-2009. We hypothesize that this is due to differences in the way the "diagonal runway", Runway 23, has been deployed.

In Figure H-1, we note that most of the fitted curves for this airport are convex, and not concave. This choice is driven by the shape of the measurements of the departure throughput as a function of the arrival throughput. For instance, Figure H-2 shows the mean value of the departure throughput at all values of arrival throughput for runway configuration 18R, 23 | 18R, 18L in 2007, and the fitted convex function. Clearly, the measured mean departure throughput is a convex function of the arrival throughput.

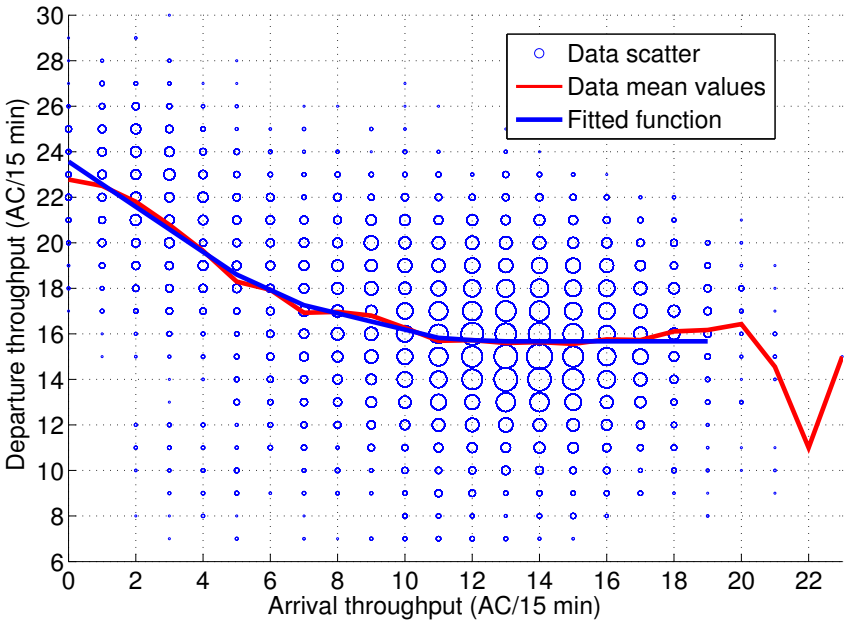


Figure H-2: Data scatter, mean values, and fitted throughput function the south flow configuration (18R, 23 | 18R, 18L) in 2007.

As Morisset demonstrates, the theoretical capacity envelopes form a convex hull of all operating points by definition [84]. If two points are valid points of the capacity envelope, like points (1, 22) and (19, 16) of the envelope of Figure H-2, any linear combination between the two points is a feasible operating point. It is achieved by operating some time at the first point and the rest of the time at the second point. For example, operating point (10,19) is simply achieved by operating half of the 15-minute period in the (1, 22) regime and the other half in the (19, 16) regime. All points between two points of the envelope are feasible and thus the capacity envelope is the convex hulls of these points. Thus, theoretically, the departure throughput is expected to be a concave function of the arrival throughput.

However, in practice this does not need be the case. As Figure H-2 shows, the departure throughput may not to be a concave function of the arrival throughput for an airport with a

complex layout. We hypothesize that this happens when the available capacity is not utilized in the same manner all the times, but it changes depending on aircraft fixes for the departure aircraft, gate location of the arrival aircraft, pilots' preferences, operating practices and other subjective factors, like congestion on the ground, delay severity etc.

For the particular example of Figure H-2, the runway configuration is most efficiently utilized in the balanced operations scenario as following: Each arrival on runway 23 is followed by a departure on runway 18L. Analysis of ASDE-X data on airports with crossing runways, where the one runway is used for arrivals and the other for departures (LGA runways 22 and 13 and BOS runways 27 and 33L) shows that 9 arrivals/15 min and 9 departures/15 min is a feasible operating point for such cases. Additionally, Runway 18R can serve six arrivals and six departures in a 15 minute interval. Thus, the balanced capacity is theoretically expected to be around 15 arrivals/15 min and 15 departures/15 min. Indeed, the estimated balanced operations capacity is 16 arrivals/15 min and 16 departures/15 min.

Similarly, runway 18R can serve two arrivals and 10 departures in a 15-minute interval. This would yield the operating point (11,19). In general, operating points (0,12), (1, 11), ..., (5,7), (6,6) are valid for Runway 18R. At the same time, Runways 23 and 18L can operate at the (9,9) point. This would yield a linear segment in the operational throughput envelope, of slope -1, between points (9,21) and (15,15). However, as Figure H-2 reveals, the operational throughput envelope stays almost flat at around 16 departures/15 min when the arrival throughput increases from 9 to 15 AC/15 min.

We hypothesize that this discrepancy results from different utilization of the runways. Consider for example the following case: On average, 5 arrivals/15 min and 7 departures/15 min are performed on Runway 18L. As arrival demand increases, arrival throughout on Runway 23 increases from 4 AC/15 min to 10 AC/15 min, and the departure throughput of Runway 18R stays flat around 9 AC/15 min. As it was shown with the example of LGA in Figure C-2, the departure throughput of a departure runway that crosses an arrival runway changes very little with the arrival throughput of the crossing runway. Thus, in this case and in agreement with Figure H-2, the airport departure throughput stays flat at 16 AC/15 min, as the arrival throughput increases from 9 to 15 AC/15 min. ASDE-X data would be helpful for investigating the utilization of the runways in more detail and for identifying opportunities for more efficient utilization of the runways.

H.2 Estimation of added capacity at CLT

Tables H.1 and H.2 summarize the capacities of CLT before and after the addition of the new runway for the two major configurations, as measured in the years 2007-2009 and 2010-2011. We observe that the balanced operations capacity increased by 25%, the arrival priority capacity by 20%, and the departure priority capacity by 29% for the south flow configuration. For the north flow configuration, the balanced operations capacity increased by 21%, the arrival priority capacity by 16%, and the departure priority capacity by 17%. Overall, from Tables H.1 and H.2, we conclude that the available capacity increased significantly.

Table H.1: CLT aggregate average runway throughput before and after the capacity expansion for the south flow configuration (AC/15 min).

South flow	2007-2009			2010-2011		
Capacity	Balanced operations	Arrival priority	Departure priority	Balanced operations	Arrival priority	Departure priority
Arrival	16	19	6	20	23	10
Departure	16	16	18	20	19	21
Total	32	35	24	40	42	31

Table H.2: CLT aggregate average runway throughput before and after the capacity expansion for the north flow configuration (AC/15 min).

North flow	2007-2009			2010-2011		
Capacity	Balanced operations	Arrival priority	Departure priority	Balanced operations	Arrival priority	Departure priority
Arrival	15	17	6	18	21	8
Departure	14	14	18	17	15	20
Total	29	31	24	35	36	28

Tables H.1 and H.2 are useful for measuring the increase in available capacity, but they do not show the actual utilization of the additional capacity. For this, we use the fitted curves of Figure H-1 and the arrival throughput measurements in the years 2009-2011. For all times in 2010 and 2011 that the north and south flow configurations were in use, the operational throughput is defined by the sum of the arrival throughput and the corresponding departure throughput (as measured with the curves "2010 Average throughput" and "2011 Average throughput" of Figure H-1). We also measure the operational throughput in the scenario of no new runway, by summing the arrival throughput at this minute and the corresponding departure throughput (as measured with the curve "2009 Average throughput").

For example, if the airport was in south flow in 2010 and the arrival throughput was 10 AC/15

min, Figure H-1a implies that the departure throughput was 21 AC/15 min. Thus, the total operational throughput at this time was 21+10=31 AC/15 min. By contrast, in 2009, for 10 arrivals/15 min, the departure throughput was only 18 AC/15 min and the total operational throughput 28 AC/ 15 min. The operational throughput improvement is 3 AC/15 min. For the times with arrival throughput larger than the arrival priority capacity in 2009 (arrival throughout $\geq$ 19 AC/15 min in Figure H-1a), the improvement is calculated by adding the improvements in arrival throughput and the departure throughput. Consider the following example: An arrival throughput of 20 AC/15 min in 2010 is higher than the arrival priority capacity of 2009. This suggests that the airport would have operated at the maximum arrival capacity of 2009, had the new runway not been added. Thus, the operational point would have been (19,16) in 2009, which implies an operational throughput of 35 AC/15 min. In 2010, arrival throughput of 20 AC/15 min implies departure throughput of 20 AC/15 min, thus the operational throughput is 40 AC/15 min and the total throughput improvement is 5 AC/15 min.

Repeating this calculation for all datapoints in the south flow configuration, we measure a total operational throughput increase of 12% in 2010, as compared to 2009. For 2011, we measure an improvement of 11%, as compared to 2009. Similarly for datapoints in the north flow configuration in 2010 and 2011, we obtain a total capacity increase of 12% compared to 2009. We also note that the departure throughput increase is the primary contributor to the increase of the operational throughput. For most of the times, the arrival throughput in 2010 and 2011 was within the arrival priority capacity limits of 2009. Thus, the new runway contributed primarily towards the departure throughput, which increased 14-16%. In the next section, we estimate the impact of this added departure throughput on the taxi-out delays.

Table H.3: CLT aggregate runway operational throughput improvement in 2010-2011, relative to the operational throughput in 2009.

	South flow			North flow		
Year	Total thr. increase	Arr. thr. increase	Dep. thr. increase	Total thr. increase	Arr. thr. increase	Dep. thr. increase
2010	12%	8%	15%	12%	5%	16%
2011	11%	7%	14%	12%	4%	16%

H.3 Taxi times at CLT

As a next step, we would like to evaluate the impact of the added capacity on the taxi-out times at CLT. In Table H.4 we list the number of movements and the average taxi times at CLT in the years 2006-2011 according to the ASPM database [38]. We notice that the traffic increased in the years 2006-2008, and the average taxi-out time also increased from 17.30 min to 18.92 min. In 2009, there were fewer movements than in 2007. However, the average taxi-out time decreased very little, and exceeded that of 2007. Clearly, the taxi-out times are a function not only of the departure demand, but also the pushback schedule, the runway configuration utilization, the weather, downstream constraints, gate locations of the departing flights etc. The average taxi-in times were stable between 6 and 6.6 min in the years 2006-2009.

In 2010, the new runway became operational, and this resulted in a significant decrease of taxi-out times despite the increase in the number of movements. In 2011, the traffic increased further and exceeded 500,000 movements, but the taxi-out times remained lower than in the years 2007-2009. We also notice the sharp increase of taxi-in times in the years 2010 and 2011. It could be conjectured that this increase is related to the addition of runway 18R/36L, which is primarily an arrival runway, and is located further from the terminal than the other runways.

Table H.4: CLT aggregate taxi times in years 2007-2011.

Year	Departures	Average taxi-out time (min)	Arrivals	Average taxi-in time (min)
2006	234,974	17.30	234,521	6.00
2007	241,960	17.95	241,616	6.59
2008	248,162	18.92	248,224	6.54
2009	238,674	18.47	238,612	6.27
2010	248,430	17.16	248,236	7.46
2011	254,188	17.85	253,995	8.54

H.4 Impact of added capacity on taxi-out times

It is clear that the aggregate taxi time statistics provided in Section H.3 do not fully describe the impact of the added runway capacity on the taxi-out times in CLT. The resultant observed taxi-out times are a complicated function of the departure demand, the pushback schedule, the arrival schedule, the added capacity, the runway configurations utilized, the gates utilized, the downstream constraints, etc. For this reason, we calibrate the queuing model described in Chapter

4 for the major runway configurations at CLT, and predict taxi-out times in 2010 and 2011 under two different scenarios, keeping everything else equal:

1. No new runway (counter-factual scenario).

2. New runway, added in 2010.

By comparing the taxi-out times and taxi-out delays predictions of the model in the two scenarios, we can evaluate the impact of the added runway capacity in the taxi-out operations. We also note that the fleet mix at CLT did not change significantly in the years 2009-2011.

H.4.1 North flow configuration

No new runway

We first train the queuing model using 2009 data, the most recent year before the addition of the new runway.

Using the filtered dataset in saturation, which was also used for deriving the operational throughput envelope shown in Figure H-1b, we construct a regression tree predicting the departure throughput as a function of the arrival throughput. We only have access to ASPM data, and do not have information on downstream constraints. The resultant (pruned) tree is shown in Figure H-3. We note that we do not have information on runway assignments either, thus the service process of the two departure runways, 36L and 36R, is modeled with a single queuing system, the parameters of which are estimated from the regression tree in Figure H-3.

We also estimate the unimpeded taxi-out times using the method described in Section 4.3.1. The average taxi-out time is 13.43 min. The taxiway congestion parameter, α is estimated at 0.15 min/AC. We then use the model parameters and the actual pushback schedule to predict taxi-out times for all flights that used runway configuration 36L, 36R | 36L, 36R in the years 2007-2009. The pushback schedules of the years 2007 and 2008 are used as test datasets. The aggregate results are summarized in Table H.5. We notice that the model predicts the average taxi-out times accurately for the training as well as the two test datasets.

Finally, we use this model to predict the taxi-out times for the flights that used runway configuration 36C, 36L, 36R | 36C, 36R in 2010 and 2011. The predicted taxi-out times and taxi-out delays are listed in Table H.5. The large difference between the actual and the predicted taxi-out times in the counterfactual scenario reflects the impact of the added capacity on the taxi-out times. In

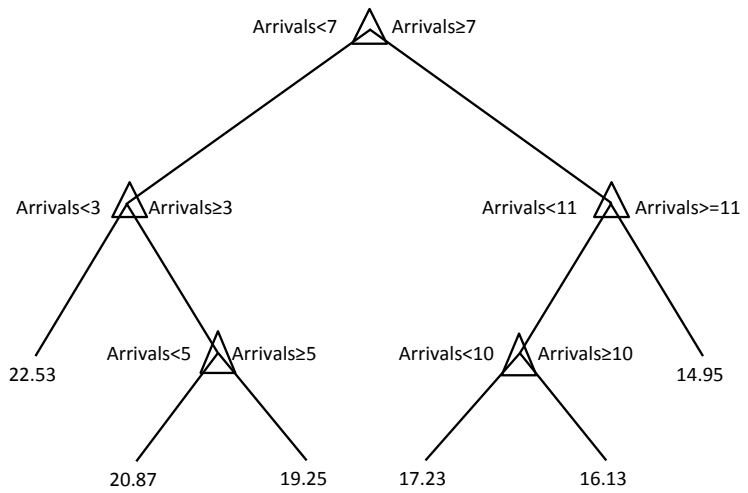


Figure H-3: Regression tree showing the departure throughput (AC/15 min) of the (VMC; 36L, 36R | 36L, 36R) configuration in 2009, parametrized by arrival throughput (AC/15 min).

2011, the mean taxi-out time would have been 23.76 min had the new runway not been built, that is, had the flights which used 36C, 36L, 36R | 36C, 36R used runway configuration 36C, 36R | 36C, 36R instead. The actual mean taxi-out time was only 19.54 min, that is 4.22 min less than what would have been in the absence of the additional runway. Similarly, taxi-out times in 2010 would have been, on average, longer by 3.76 min for the 40,497 flights that used runway configuration 36C, 36L, 36R | 36C, 36R, had the new runway not been constructed.

Table H.5: CLT aggregate taxi-out time predictions for the north flow configuration during the years 2007-2011. No new runway is assumed in the model predictions.

Year	Departures	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Model mean taxi-out delay (min)	Actual med. taxi-out time (min)	Model med. taxi-out time (min)
2007	74,216	19.81	19.63	6.20	18	18
2008	64,429	21.60	21.13	7.70	19	20
2009	59,631	22.37	21.93	8.50	20	20
2010	40,497	19.46	23.22	9.79	18	22
2011	43,005	19.54	23.76	10.33	18	22

New runway

We first train the queuing model using 2011 data, the most recent year after the addition of the new runway. Using the filtered dataset in saturation, which was also used for deriving the operational throughput envelope shown in Figure H-1b, we construct a regression tree predicting the departure throughput as a function of the arrival throughput. The resultant (pruned) tree is shown in Figure

H-4. Comparing it to the regression tree shown in Figure H-3, we immediately notice the increase in departure throughput under all arrival throughput conditions.

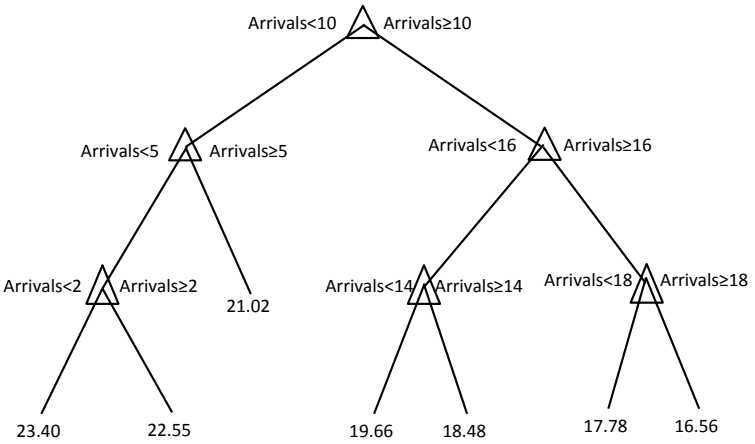


Figure H-4: Regression tree showing the departure throughput (AC/15 min) of (VMC; 36C, 36L, 36R | 36C, 36R) configuration in 2011 parametrized by arrival throughput (AC/15 min).

We also estimate the unimpeded taxi-out times as described in Section 4.3.1. The average taxi-out time is 13.28 min, very similar to that of 2009. This similarity was expected because the new runway is not used for departures in this runway configuration. The taxiway congestion parameter, α is estimated at 0.13 min/AC. We then use the model parameters and the actual pushback schedule to predict taxi-out times for all flights that used runway configuration 36C, 36L, 36R | 36C, 36R in the years 2010 and 2011. The pushback schedule for 2010 is used as the test dataset. The aggregate results are summarized in Table H.6. We notice that the model predicts the average taxi-out times accurately for the training and the test dataset.

Table H.6: CLT aggregate taxi-out time predictions for the north flow configuration for the years 2010-2011.

Year	Departures	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Model mean taxi-out delay (min)	Actual med. taxi-out time (min)	Model med. taxi-out time (min)
2010	40,497	19.46	19.97	6.70	18	19
2011	43,005	19.54	19.24	5.96	18	18

Predicted delay reduction

We now compare the predicted taxi-out times and the predicted delays of the two scenarios and estimate the taxi-out delay reduction resulting from the added capacity in Table H.7. From Tables

H.3 and H.7, we notice that as predicted by the model, the 16% increase in departure operational throughput resulted in 32% and 42% decrease in taxi-out delays tin 2010 and 2011 respectively.

Table H.7: CLT predicted taxi-out time and taxi-out delay reduction for the north flow configuration in 2010 and 2011

Year	Departures	Predicted mean taxi-out time reduction (min)	Predicted mean taxi-out delay reduction (min)	Predicted percent taxi-out time reduction	Predicted percent taxi-out delay reduction
2010	40,497	3.25	3.09	14%	32%
2011	43,005	4.53	4.37	19%	42%

These results can also be visualized using the figures introduced in Chapters 4 and 5. Figure H-5 shows the frequency of the different congestion states observed in the operational data and predicted by the model in 2010 (the test dataset). The model predicts the frequency of congestion levels reasonably well. Figure H-5 also shows the expected taxi-out time as a function of the number of aircraft taxiing-out at the time of pushback, for both actual operations and modeled operations. Figure H-6 shows the predicted throughput of configuration 36C, 36L, 36R | 36C, 36R at CLT in 2010 as a function of the congestion state N. We note that the model predicts the mean and the median throughput and the mean taxi-out times very accurately in all traffic conditions, despite modeling the two departure runways as a single server, and despite the slightly different airport capacity in 2010 (Figure H-1b).

These results can be contrasted with the results shown in Figures H-7 and H-8, which show the corresponding predictions in the counter-factual scenario, that is, had the new runway not been added. The benefit of the added departure capacity in terms of reducing both the congestion and the taxi-out times can be clearly seen in these figures.

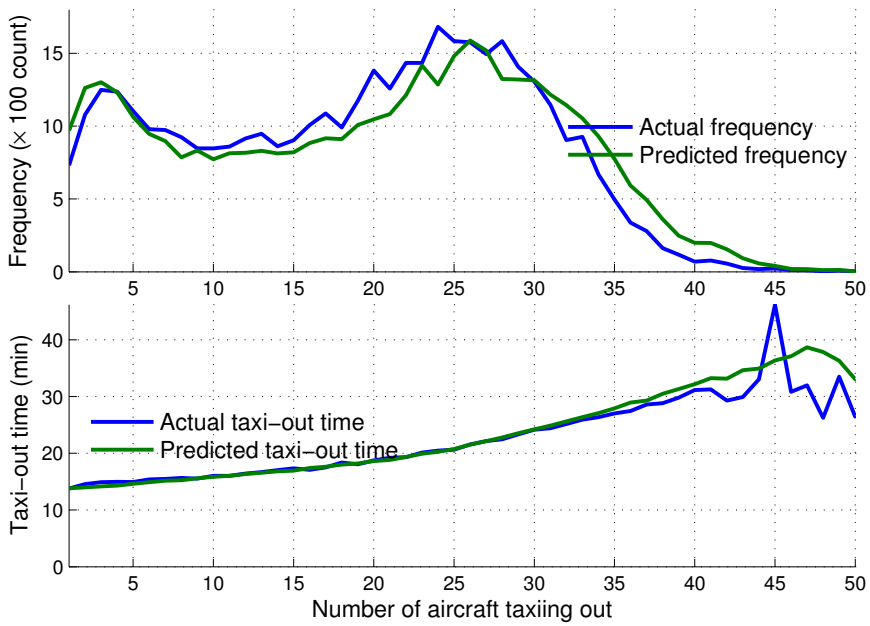


Figure H-5: Actual and modeled frequency of each state N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for CLT runway configuration 36C, 36L, 36R | 36C, 36R in 2010.

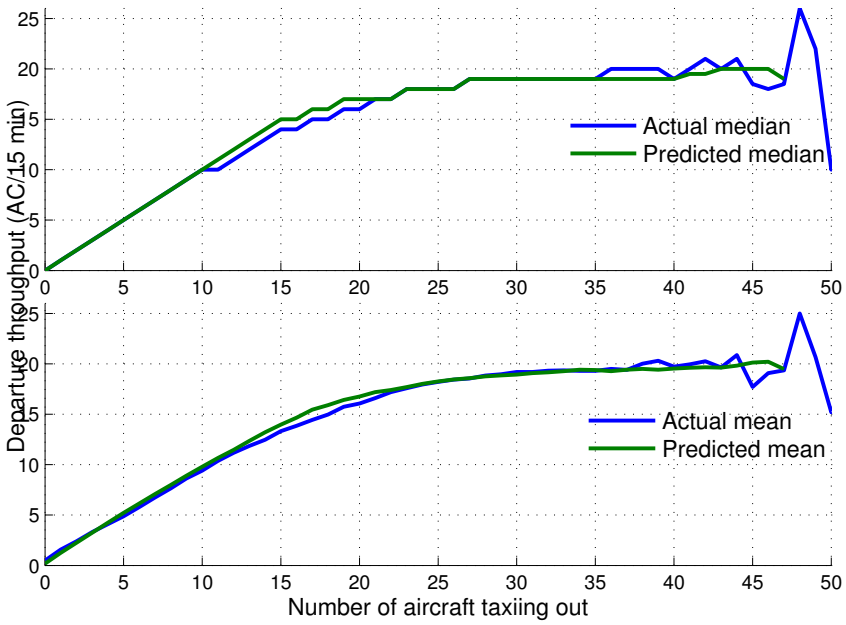


Figure H-6: Actual and modeled throughput of all states N for CLT runway configuration 36C, 36L, 36R | 36C, 36R in 2010: Mean (top); Median (bottom).

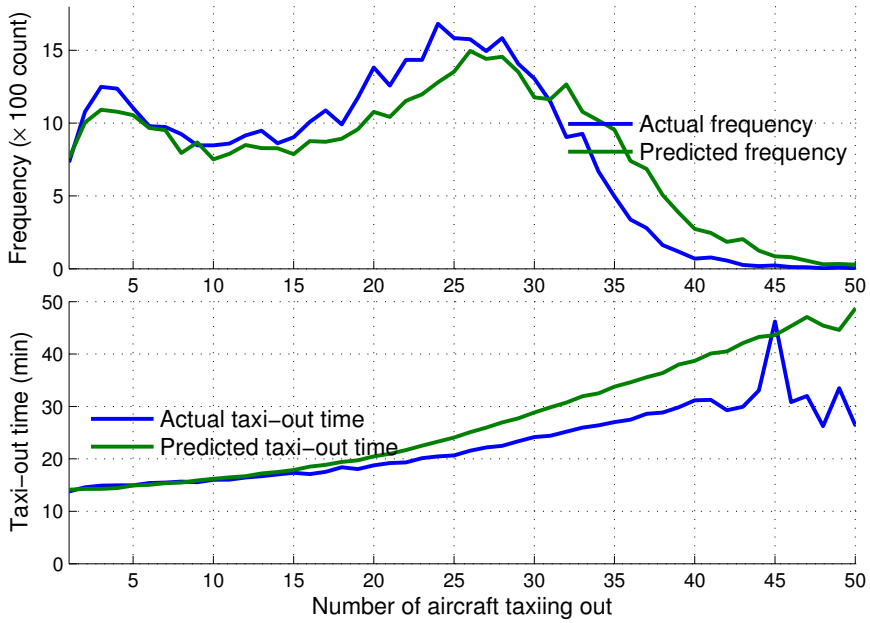


Figure H-7: Actual and modeled frequency of each state N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for CLT runway configuration 36C, 36R | 36C, 36R in 2010. No new runway is assumed in the model predictions.

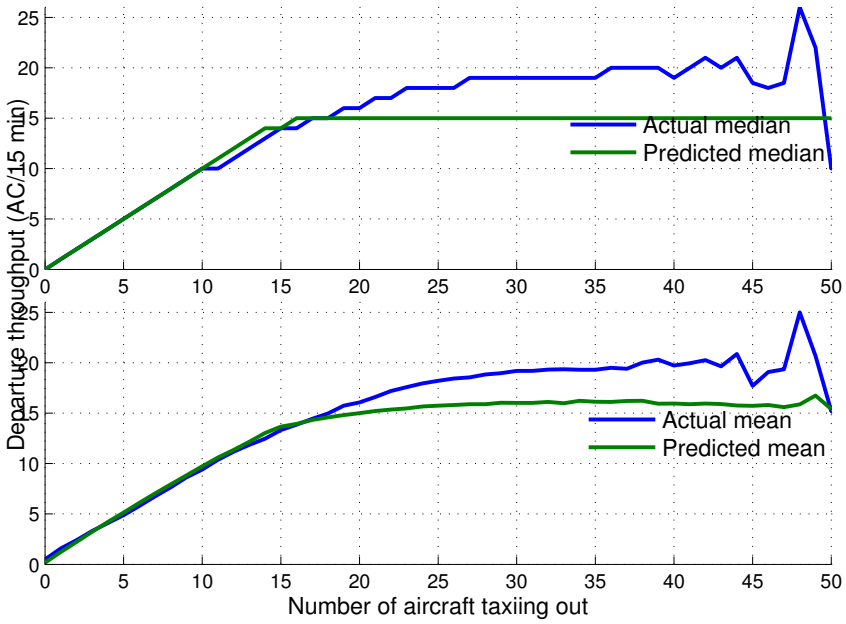


Figure H-8: Actual and modeled throughput of all states N for CLT runway configuration 36C, 36R | 36C, 36R in 2010: Mean (top); Median (bottom). No new runway is assumed in the model predictions.

We also use the model to predict the average throughput and taxi-out times evolution during an average day at CLT in 2010. In the upper plot of Figure H-9, we show the average number of takeoffs (or departures) that was recorded during each 15 minute time window for all days in which this runway configuration was in use in 2010. We also show the average number of departures of this runway configuration in CLT, as predicted by the model. In the lower part of Figure H-9, we show the actual and predicted average taxi-out times for the flights that pushed back in each 15-minute time window. We observe that the model faithfully represents an average day at CLT.

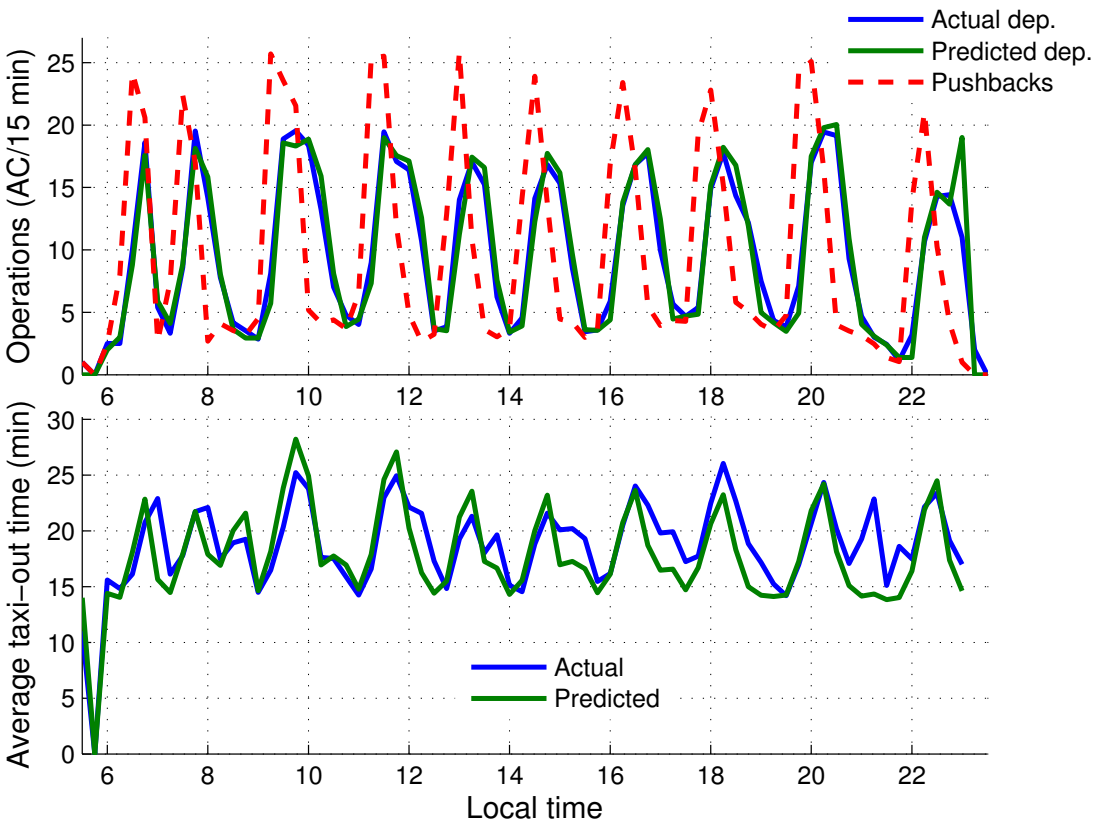


Figure H-9: Average number of pushbacks, and average numbers of actual and predicted takeoffs by time of day at CLT for runway configuration 36C, 36L, 36R | 36C, 36R in 2010 (top); Average actual and predicted taxi-out times (bottom).

Analogously, we show the predictions had the new runway capacity not been added in Figure H-10. We notice that the absence of the new runway leads to much longer taxi-out times during most departure pushes. Comparing Figures H-9 and H-10, we notice that the added capacity especially benefits the flights that would have suffered the longest delays. Without the new runway, flights that push back between 1000 hours and 1015 hours are predicted to have taxi-out times of 38 min on average. With the new runway, the actual and predicted taxi-out times for these flights are only

26-27 minutes.

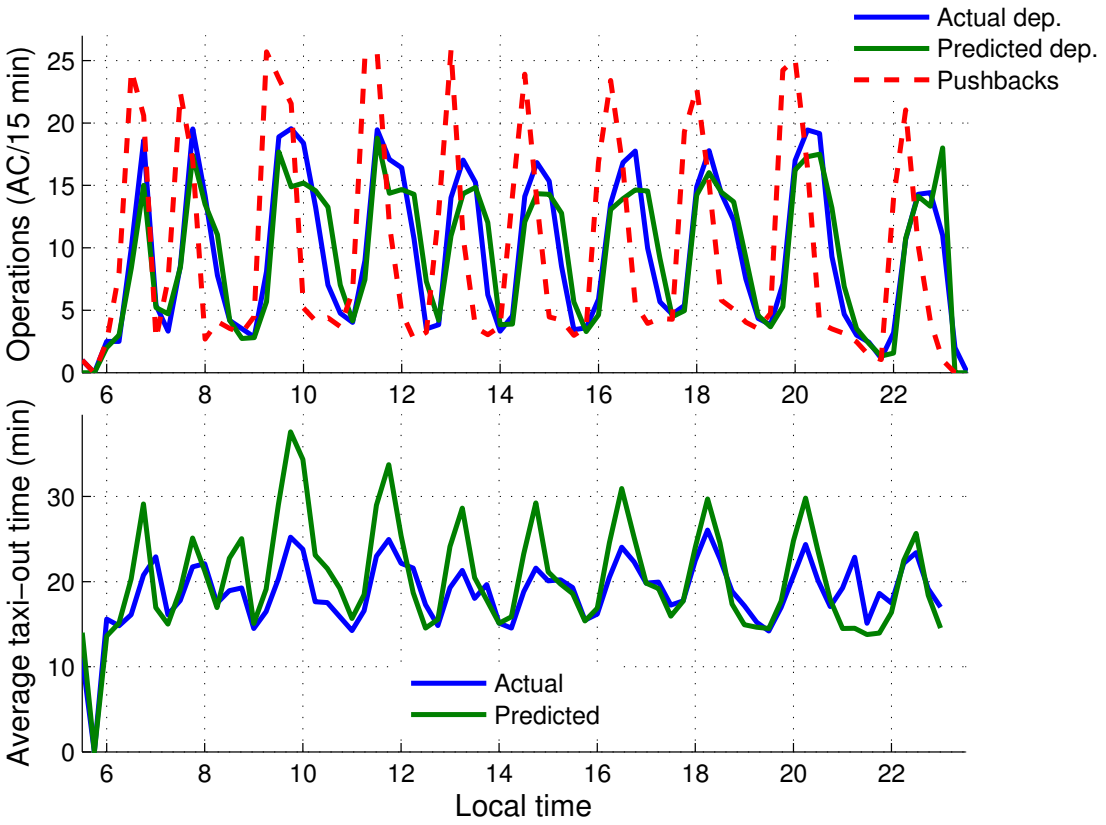


Figure H-10: Average number of pushbacks, and average numbers of actual and predicted takeoffs by time of day at CLT for runway configuration 36C, 36R | 36C, 36R in 2010 (top); Average actual and predicted taxi-out times (bottom). No new runway is assumed in the model predictions.

H.4.2 South flow configuration

Here, we repeat the process described in the previous section for the south flow runway configuration. We first train the queuing model using 2009 data, the most recent year before the addition of the new runway.

We use the model parameters and the actual pushback schedule to predict taxi-out times for all flights that used runway configuration 18R, 23 | 18R, 18L in the years 2007-2009. The pushback schedules of the years 2007 and 2008 are used as test datasets. The aggregate results are summarized in Table H.8. We notice that the model does not predict the average taxi-out times very accurately for the test datasets. This is because of the differences in the operational throughput curves in the years 2007-2009 shown in Figure H-1a, and discussed in Section H.1. Clearly, the queuing model is not capable of predicting changes in the airport capacity across different years.

Subsequently, we use the model to predict the taxi-out times for the flights that used runway configuration 18C, 18R, 23 | 18C, 18L in in 2010 and 2011. The predicted taxi-out times and taxi-out delays are listed in Table H.8 as well. As in the case of the north flow configuration, the significant difference between the actual and the predicted taxi-out times in the counterfactual scenario indicates the impact of the added capacity on the taxi-out times. In 2011, the mean taxi-out time would have been 19.64 min had the new runway not been built, that is, had the 61,540 flights which used runway 18C, 18R, 23 | 18R, 18L used runway configuration 18C, 23 | 18C, 18L instead. The actual mean taxi-out time was only 16.73 min, that is 2.91 min less than what would have been in the absence of the additional runway. Similarly, taxi-out times in 2010 would have been on average 2.71 min longer for the 52,965 flights that used runway configuration 18C, 18R, 23 | 18C, 18L, had the new runway not been constructed.

Table H.8: CLT aggregate taxi-out time predictions for the south flow configuration in years 2007-2011. No new runway is assumed in the model predictions.

Year	Departures	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Model mean taxi-out delay (min)	Actual med. taxi-out time (min)	Model med. taxi-out time (min)
2007	122,879	17.11	16.06	4.01	15	15
2008	104,233	18.08	16.85	4.80	16	16
2009	108,843	18.55	18.14	6.10	17	17
2010	52,965	16.56	19.27	7.23	15	18
2011	61,540	16.73	19.64	7.59	15	19

We then train the queuing model using 2011 data, the most recent year after the addition of the new runway. We use the model parameters and the actual pushback schedule to predict taxi-out times for all flights that used runway configuration 18C, 18R, 23 | 18C, 18L in the years 2010 and 2011. The pushback schedules for 2010 is used as the test dataset. The aggregate results are summarized in Table H.9. We notice that the model predicts the average taxi-out times fairly well for the training and the test datasets.

Table H.9: CLT aggregate taxi-out time predictions for the south flow configuration for the years 2010-2011.

Year	Departures	Actual mean taxi-out time (min)	Model mean taxi-out time (min)	Model mean taxi-out time (min)	Actual med. taxi-out time (min)	Model med. taxi-out time (min)
2010	52,965	16.56	16.78	5.26	15	16
2011	61,540	16.73	15.93	4.42	15	15

We compare the predicted taxi-out times and the predicted delays of the two scenarios, and estimate the taxi-out delay reduction resulting from the added capacity in Table H.10. From Tables H.3 and H.10, we notice that as predicted, the 15% and 14% increase in departure operational throughput in 2010 and 2011 resulted in 27% and 42% decreases in taxi-out delays, respectively.

Table H.10: CLT predicted taxi-out time and taxi-out delay reduction for the south flow configuration in 2010 and 2011.

Year	Departures	Predicted mean taxi-out time reduction (min)	Predicted mean taxi-out delay reduction (min)	Predicted percent taxi-out time reduction	Predicted percent taxi-out delay reduction
2010	52,965	2.49	1.97	13%	27%
2011	61,540	3.70	3.18	19%	42%

Finally, we provide visualizations of the ability of the model to predict operations in a multi-runway system with crossing runways by using a single stochastic and dynamic queuing system in Figures H-11 and H-12, which show the predictions for 2010 (the test dataset). In Figure H-12, we notice that the model predicts the average departure throughput at all congestion states very accurately. Similarly, from Figure H-11, we notice that it predicts taxi-out times very accurately at lower and medium congestion states ($N < 30$).

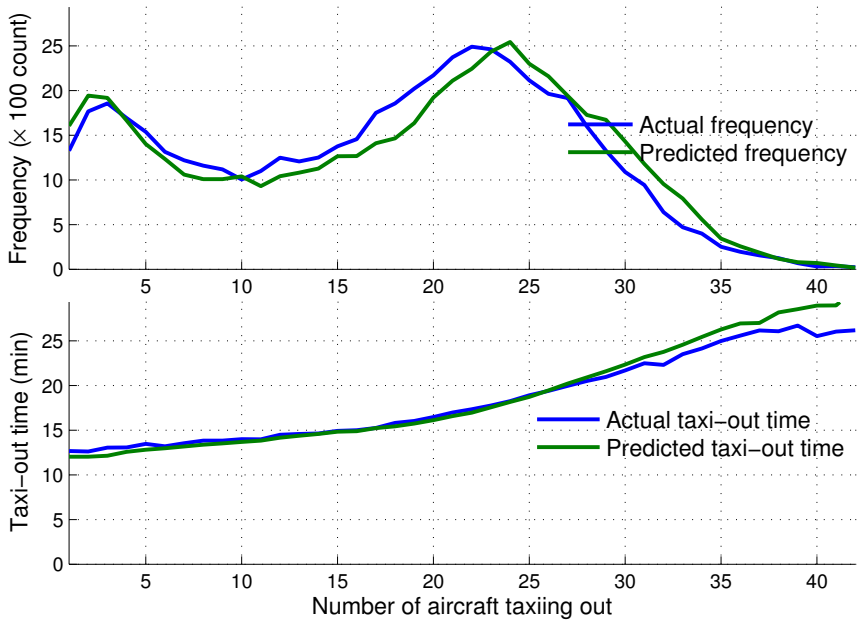


Figure H-11: Actual and modeled frequency of each state N (top); Actual and modeled average taxi-out time as a function of the state N at the time of pushback (bottom) for CLT runway configuration 18C, 18R, 23 | 18C, 18L in 2010.

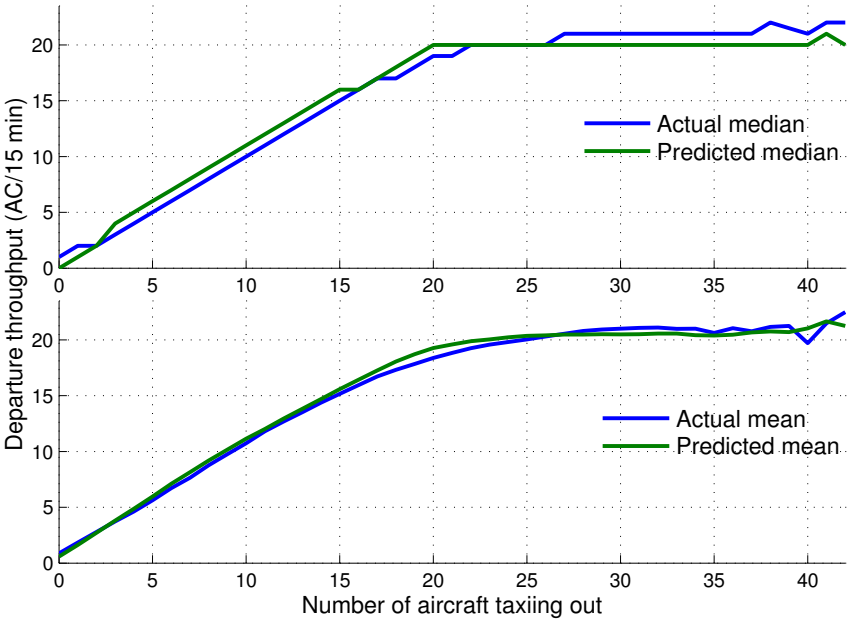


Figure H-12: Actual and modeled throughput of all states N for CLT runway configuration18C, 18R, 23 | 18C, 18L in 2010: Mean (top); Median (bottom).

H.5 Conclusions

In this appendix, we assessed the addition of a new runway at CLT in 2010. Firstly, we found that the operational throughput envelopes for a multi-runway airport, where runways are shared by arrivals and departures, are not necessarily convex. We measured the capacities of the two major runway configurations of the airport before and after the addition of the runway. We found that the balanced operations capacity of the south flow increased from 32 AC/15 min to 40 AC/15 min, and that of the north flow increased from 29 AC/15 min to 35 AC/15 min. We also found that the operational departure throughput increased on average by only 14% to 16% for the two major runway configurations respectively. Finally, we calibrated the queuing model for the north flow and the south flow runway configurations, and predicted the taxi-out times and the taxi-out delays with and without the new runway. We estimated that taxi-out delays decreased approximately by 30% in 2010 and by 40% in 2011, as a result of added capacity.

Bibliography

[1] I. Anagnostakis and J.P. Clarke. Runway operations planning: A two-stage heuristic algorithm. In *Aviation Technology, Integration, and Operations Conference (ATIO)*, 2002.

[2] I. Anagnostakis and J.P. Clarke. Runway operations planning: A two-stage solution methodology. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, page 12. IEEE, 2003.

[3] I. Anagnostakis, J.P. Clarke, D. Bohme, and U. Volckers. Runway operations planning and control: sequencing and scheduling. *Journal of Aircraft*, 38:988–997, 2001.

[4] S. Atkins and D. Walton. Prediction and control of departure runway balancing at dallas/fort worth airport. In *American Control Conference, 2002. Proceedings of the 2002*, volume 2, pages 1308–1313. IEEE, 2002.

[5] H. Balakrishnan and B. Chandran. Efficient and Equitable Departure Scheduling in Real-time: New Approaches to Old Problems. In *7th USA-Europe Air Traffic Management Research and Development Seminar*, 2007.

[6] H. Balakrishnan and B. Chandran. Algorithms for Scheduling Runway Operations under Constrained Position Shifting. *Operations Research*, 58(6):1650, 2010.

[7] H. Balakrishnan and Y. Jung. A Framework for Coordinated Surface Operations Planning at Dallas Fort Worth International Airport. In *AIAA Guidance, Navigation, and Control Conference*, Hilton Head, SC, August 20-23 2007.

[8] M. Ball, C. Barnhart, G. Nemhauser, A. Odoni, C. Barnhart, and G. Laporte. Air transportation: Irregular operations and control. *Handbooks in Operations Research and Management Science*, 14(1):1–67, 2007.

[9] M. Ball, M. Hansen, P. Swaroop, and P. Zou. Do more U.S. airports need slot controls? A welfare based approach to determine slot levels, 2011.

[10] JE Beasley, M. Krishnamoorthy, YM Sharaiha, and D. Abramson. Scheduling aircraft landings-the static case. *Transportation Science*, 34(2):180–197, 2000.

[11] D.P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, 2005.

[12] D. Bertsimas and S. S. Patterson. The air traffic flow management problem with enroute capacities. *Operations Research*, 46(3):406–422, 1998.

[13] D. Böhme. Tactical departure management with the Eurocontrol/DLR DMAN. In *6th USA/Europe Air Traffic Management Research and Development Seminar, Baltimore, MD*, 2005.

[14] C. Brinton, C. Provan, S. Lent, T. Prevost, and S. Passmore. Collaborative Departure Queue Management: An Example of Collaborative Decision Making in the United States. In *9th USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Berlin, Germany, June 2011.

[15] P. Burgain. *On the control of Airport Departure Processes*. PhD thesis, Georgia Institute of Technology, 2010.

[16] P. Burgain, E. Feron, J.P. Clarke, and A. Darrasse. Collaborative Virtual Queue: Fair Management of Congested Departure Operations and Benefit Analysis. *Arxiv preprint arXiv:0807.0661*, 2008.

[17] P. Burgain, O.J. Pinon, E. Feron, J.P. Clarke, and DN Mavris. On the value of information within a collaborative decision making framework for airport departure operations. In *Digital Avionics Systems Conference*. IEEE, 2009.

[18] F. Carr, A. Evans, J.P. Clarke, and E. Feron. Modeling and control of airport queueing dynamics under severe flow restrictions. In *Proceedings of the American Control Conference*, volume 2, 2002.

[19] F. Carr, A. Evans, E. Feron, and J.P. Clarke. Software tools to support research on airport departure planning. In *Digital Avionics Systems Conference*, Irvine CA, 2002. IEEE.

[20] F.R Carr. Stochastic modeling and control of airport surface traffic. Master's thesis, Massachusetts Institute of Technology, 2001.

[21] D. C. Carslaw, S. D. Beevers, K. Ropkins, and M. C. Bell. Detecting and quantifying aircraft and other on-airport contributions to ambient nitrogen oxides in the vicinity of a large international airport. *Atmospheric Environment*, 40(28):5424–5434, 2006.

[22] B. Chandran and H. Balakrishnan. A dynamic programming algorithm for robust runway scheduling. In *American Control Conference*. IEEE, 2007.

[23] J.P. Clarke. Surface traffic optimization in the presence of uncertainties. Technical report, 2010. Technical Report NASA/Contract No. NNX07AU34A.

[24] J.P. Clarke, T. Melconian, E. Bly, and F. Rabbani. MEANS MIT Extensible Air Network Simulation. *Simulation*, 83(5):385, 2007.

[25] R. Clewlow and D. Michalek. Logan Control Tower: Controller Positions, Processes, and Decision Support Systems. Technical report, Massachusetts Institute of Technology, 2010.

[26] R.R. Clewlow, I. Simaiakis, and H. Balakrishnan. Impact of Arrivals on Departure Taxi Operations at Airports. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2010.

[27] J. Cohen. *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum, 2003.

[28] Conductive Technology. Flightstats. `http://www.flightstats.com`, accessed May 2012.

[29] T.B. Crabill, D. Gross, and M.J. Magazine. A classified bibliography of research on optimal design and control of queues. *Operations Research*, 25(2):219–232, 1977.

[30] Richard de Neufville and Amedeo Odoni. *Airport Systems: Planning, Design and Management*. McGraw-Hill, 2003.

[31] R. DeLaura and S. Allan. Route selection decision support in convective weather: a case study of the effects of weather and operational assumptions on departure throughput. In *5th Eurocontrol/FAA ATM R&D Seminar, Budapest, 23–27 June 2003*, 2003.

[32] R. DeLaura, M. Robinson, R. Todd, and K. MacKenzie. Evaluation of weather impact models in departure management decision support: operational performance of the Route Availability Planning Tool (RAPT) prototype. In *13th Conference on Aviation, Range, and Aerospace Meteorology, AMS, New Orleans, LA*, 2008.

[33] A. D. Donaldson and R.J. Hansman. Improvement of terminal area capacity in the New York airspace. Technical report, Massachusetts Institute of Technology, 2011. No. ICAT-2011-4.

[34] Heinz Erzberger. Design Principles and Algorithms for Automated Air Traffic Management. Technical report, NASA, 1995.

[35] T.P. Fan and A.R. Odoni. The potential of demand management as a short-term means of relieving airport congestion. In *Proceedings of EUROCONTROL-FAA Air Traffic Management R&D Review Seminar, Santa Fe, NM*, 2001.

[36] Federal Aviation Administration. Operating Limitations at Newark Liberty International Airport; Order Limiting Scheduled Operations at Newark Liberty International Airport. *Federal Register 73(99) 29550, 29566*, 2008.

[37] Federal Aviation Administration. Fact Sheet of Airport Surface Detection Equipment, Model X (ASDE-X), October 2010.

[38] Federal Aviation Administration. Aviation System Performance Metrics database. `http://aspm.faa.gov/aspm/ASPMframe.asp`, accessed February 2012.

[39] Federal Aviation Administration. BOS Airport Acceptance Rate . `http://www.fly.faa.gov/Information/east/zbw/bos/bos_aar1.htm`, accessed February 2012.

[40] Federal Aviation Administration. CLT Airport Acceptance Rate. `http://www.fly.faa.gov/Information/east/ztl/clt/clt_aar.htm`, accessed February 2012.

[41] Federal Aviation Administration. EWR Traffic Management Tips. `http://www.fly.faa.gov/Information/east/zny/ewr/ewr_tm.htm`, accessed February 2012.

[42] Federal Aviation Administration. Terminal Procedures Publication/Airport Diagrams. `http://aeronav.faa.gov/index.asp?xml=aeronav/applications`, accessed January 2012.

[43] Federal Aviation Administration. The Operations Network. `http://aspm.faa.gov/OPSNET`, accessed March 2012.

[44] Federal Aviation Administration. Airline Service Quality Performance. `http://aspm.faa.gov/ASQP`, accessed September 2008.

[45] Federal Aviation Administration. Enhanced Traffic Management System. `http://www.fly.faa.gov/Products/Information/ETMS/etms.html`, accessed September 2008.

[46] Federal Aviation Administration Office of Aviation Policy and Plans. Documentation for the Aviation System Performance Metrics: Actual versus Scheduled Metrics. Technical report, 2002.

[47] E. R. Feron, R. J. Hansman, A. R. Odoni, R. B. Cots, B. Delcaire, W. D. Hall, H. R. Idris, A. Muharremoglu, and N. Pujet. The Departure Planner: A conceptual discussion. Technical report, Massachusetts Institute of Technology, 1997.

[48] S.B. Gershwin. Design and operation of manufacturing systems: the control-point policy. *IIE Transactions*, 32(10):891–906, 2000.

[49] E.P. Gilbo. Airport capacity: Representation, estimation, optimization. *IEEE Transactions on Control Systems Technology*, 1(3):144–154, 1993.

[50] J.B. Gotteland, N. Durand, J.M. Alliot, and E. Page. Aircraft ground traffic optimization. In *Proceedings of the 4th International Air Traffic Management R&D Seminar*, 2001.

[51] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. `http://cvxr.com/cvx`, April 2011.

[52] K.J. Griffin, P. Yu, Rappaport D.B., and J.P. Clarke. Evaluating Surface Optimization Techniques Using a Fast-time Airport Surface Simulation. In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, September 2010.

[53] S. Gupta. Transient analysis of D(t)/M(t)/1 queuing system with applications to computing airport delays. Master's thesis, Massachusetts Institute of Technology, 2010.

[54] M. Hansen, T. Nikoleris, D. Lovell, K. Vlachou, and A. Odoni. Use of Queuing Models to Estimate Delay Savings from 4D Trajectory Precision. In *Eighth USA/Europe Air Traffic Management Research and Development Seminar*, 2009.

[55] J.E. Hebert and D.C. Dietz. Modeling and analysis of an airport departure process. *Journal of Aircraft*, 34(1):43–47, 1997.

[56] G. Hengsbach and A.R. Odoni. Time dependent estimates of delays and delay costs at major airports. Technical report, Cambridge, Mass.: MIT, Dept. of Aeronautics & Astronautics, Flight Transportation Laboratory, 1975.

[57] F.S. Hillier and G.J. Lieberman. *Introduction to Operations Research*. McGraw-Hill, 2001.

[58] H. Idris, J. P. Clarke, R. Bhuva, and L. Kang. Queuing Model for Taxi-Out Time Estimation. *Air Traffic Control Quarterly*, 2001.

[59] H. R. Idris. *Observation and analysis of departure operations at Boston Logan International Airport*. PhD thesis, Massachusetts Institute of Technology, 2001.

[60] H. R. Idris, I. Anagnostakis, B. Delcaire, W. D. Hall, J.-P. Clarke, R. J. Hansman, E. Feron, and A. R. Odoni. Observations of Departure Processes At Logan Airport to Support the Development of Departure Planning Tools. *Air Traffic Control Quarterly*, 7(4), 1999.

[61] H. R. Idris, B. Delcaire, I. Anagnostakis, W. D. Hall, N. Pujet, E. Feron, R. J. Hansman, J. P. Clarke, and A. R. Odoni. Identification of flow constraint and control points in departure operations at airport systems. In *AIAA Guidance, Navigation and Control Conference*, August 1998.

[62] A. Jacquillat and A. R. Odoni. A Study of Airport Congestion at JFK and EWR. In *5th International Conference on Research in Air Transportation*, 2012.

[63] Alexandre Jacquillat. A queuing model of airport congestion and policy implications at JFK and EWR. Master's thesis, Massachusetts Institute of Technology, 2012.

[64] M. Janic. Modelling the capacity of closely-spaced parallel runways using innovative approach procedures. *Transportation Research Part C: Emerging Technologies*, 16(6):704–730, 2008.

[65] Joint Economic Committee. Flight Delays Cost Passengers, Airlines, and the U.S. Economy Billions, 2008.

[66] Y. Jung, T. Hoang, J. Montoya, G. Gupta, W. Malik, and L. Tobias. Performance Evaluation of a Surface Traffic Management Tool for Dallas/Fort Worth International Airport. In *9th USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Berlin, Germany, June 2011.

[67] 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. In *9th USA/Europe ATM Research & Development Seminar*, 2011.

[68] H. Khadilkar and H. Balakrishnan. A Multi-Modal Unscented Kalman Filter for Inference of Aircraft Position and Taxi Mode from Surface Surveillance Data. In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, Virginia Beach, VA, September 2011.

[69] H. Khadilkar, H. Balakrishnan, and B. Reilly. Analysis of airport performance using surface surveillance data: A case study of BOS. In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, Virginia Beach, VA, September 2011.

[70] H.D. Khadilkar. Analysis and modeling of airport surface operations. Master's thesis, Massachusetts Institute of Technology, 2011.

[71] P. A. Kivestu. Alternative methods of investigating the time dependent M/G/k queue. Master's thesis, Massachusetts Institute of Technology, 1976.

[72] R. Koenker and K.F. Hallock. Quantile regression. *Journal of Economic Perspectives*, pages 143–156, 2001.

[73] B.O. Koopman. Air-terminal queues under time-dependent conditions. *Operations Research*, 20(6):1089–1114, 1972.

[74] J. Le Ny and H. Balakrishnan. Feedback Control of the National Airspace System to Mitigate Weather Disruptions. In *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010.

[75] H. Lee and H. Balakrishnan. Optimization of Airport Taxiway Operations at Detroit Metropolitan Airport (DTW). In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, September 2010.

[76] H. Lee, I. Simaiakis, and H. Balakrishnan. A Comparison of Aircraft Trajectory-Based and Aggregate Queue-Based Control of Airport Taxi Processes. In *Digital Avionics Systems Conference*. IEEE, 2010.

[77] J. Legge and B. Levy. Departure taxi time predictions using ASDE-X surveillance data. In *26th International Congress of the Aeronautical Sciences*, 2008.

[78] S.A. Lippman. Applying a new device in the optimization of exponential queuing systems. *Operations Research*, 23(4):687–710, 1975.

[79] D. Long, D. Lee, J. Johnson, E. Gaier, and P. Kostiuk. Modeling Air Traffic Management Technologies with a Queuing Network Model of the National Airspace System. Technical report, NASA Langley Research Center, Hampton, VA, 1999. Technical Report NASA/CR-1999-208988.

[80] D.W. Low. Optimal dynamic pricing policies for an M/M/s queue. *Operations Research*, 22(3):545–561, 1974.

[81] K. M. Malone. *Dynamic queueing systems : behavior and approximations for individual queues and for networks*. PhD thesis, Massachusetts Institute of Technology, 1995.

[82] M. A. Miracolo, C. J. Hennigan, M. Ranjan, N. T. Nguyen, T. D. Gordon, E. M. Lipsky, A. A. Presto, N. M. Donahue, and A. L. Robinson. Secondary aerosol formation from photochemical aging of aircraft exhaust in a smog chamber. *Atmospheric Chemistry and Physics*, 11(9):4025–4610, 2011.

[83] MITRE Corporation and Federal Aviation Administration. Airport Capacity Benchmark Report 2004. Technical report, 2004.

[84] T. Morisset. Comparing Capacities and Delays at Major European and American Airports. Master's thesis, Massachusetts Institute of Technology, 2010.

[85] A. Nakahara. Estimating current and future benefits of airport surface congestion management techniques. Master's thesis, Massachusetts Institute of Technology, 2012.

[86] A. Nakahara, T. Reynolds, T. White, and R. Dunsky. Analysis of a Surface Congestion Management Technique at New York JFK Airport . In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, Virginia Beach, VA, September 2011.

[87] A.R. Odoni, J. Bowman, D. Delahaye, J.J. Deyst, E. Feron, R.J. Hansman, K. Khan, J.K. Kuchar, N. Pujet, and R.W. Simpson. Existing and required modeling capabilities for evaluating ATM systems and concepts. Technical report, 1997.

[88] Office of Aviation Policy and Plans, Federal Aviation Administration. Documentation for the Aviation System Performance Metrics (ASPM), May 2002.

[89] M.D. Peterson, D.J. Bertsimas, and A.R. Odoni. Models and algorithms for transient queueing congestion at airports. *Management Science*, pages 1279–1295, 1995.

[90] D.M. Pfeil. *Optimization of airport terminal-area air traffic operations under uncertain weather conditions*. PhD thesis, Massachusetts Institute of Technology, 2011.

[91] N. Pujet. *Modeling and control of the departure process of conjested airports*. PhD thesis, Massachusetts Institute of Technology, 1999.

[92] N. Pujet, B. Delcaire, and E. Feron. Input-output modeling and control of the departure process of congested airports. *AIAA Guidance, Navigation, and Control Conference and Exhibit, Portland, OR*, pages 1835–1852, 1999.

[93] N. Pyrgiotis. *A Stochastic and Dynamic Model of Delay Propagation Within an Airport Network For Policy Analysis.* PhD thesis, Massachusetts Institute of Technology, 2011.

[94] N. Pyrgiotis, K.M. Malone, and A. Odoni. Modelling delay propagation within an airport network. *Transportation Research Part C: Emerging Technologies*, 2011.

[95] R. Clewlow. A Prediction Model for Aircraft Surface Delays. Technical report, MIT, 2010.

[96] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-07-0.

[97] V. Ramanujam and H. Balakrishnan. Estimation of Arrival-Departure Capacity Tradeoffs in Multi-Airport Systems. In *Proceedings of the 48th IEEE Conference on Decision Control*, 2009.

[98] S. Rathinam, J. Montoya, and Y. Jung. An Optimization Model for Reducing Aircraft Taxi Times at the Dallas Fort Worth International Airport. In *26th International Congress of the Aeronautical Sciences*, Anchorage, AK, September 16 2008.

[99] T. G. Reynolds and L. Sandberg. PHL Site Visit Report. Technical report, Massachusetts Institute of Technology, 2012.

[100] M. Robinson, R. DeLaura, J. Evans, and S. McGettigan. Operational usage of the route availability planning tool during the 2007 convective weather season. In *American Meteorological Society 13th Conference on Aviation, Range and Aerospace Meteorology*, 2008.

[101] S. Roy, B. Sridhar, and G.C. Verghese. An aggregate dynamic stochastic model for an air traffic system. In *Proceedings of the 5th USA/Europe ATM 2003 R&D Seminar*, 2003.

[102] L. Sandberg. Applications of ASDE-X Data to the Analysis of Airport Surface Operations. Master's thesis, Massachusetts Institute of Technology, 2012.

[103] M. Schaper, G. Tsoukala, R. Stavrati, and N. Papadopoulos. Departure flow control through takeoff sequence optimisation: Setup and results of trials at Athens airport. In *2011 IEEE/AIAA 30th Digital Avionics Systems Conference (DASC)*, pages 2B2–1. IEEE, 2011.

[104] R. A. Shumsky. *Dynamic statistical models for the prediction of aircraft take-off times.* PhD thesis, Massachusetts Institute of Technology, 1995.

[105] R.A. Shumsky. Real-Time Forecasts of Aircraft Departure Queues. *Air Traffic Control Quarterly*, 5(4), 1997.

[106] I. Simaiakis. Modeling and Control of Airport Departure Processes for Emissions Reduction. Master's thesis, Massachusetts Institute of Technology, 2009.

[107] I. Simaiakis and H. Balakrishnan. Queuing Models of Airport Departure Processes for Emissions Reduction. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2009.

[108] I. Simaiakis and H. Balakrishnan. Analysis and Control of Airport Departure Processes to Mitigate Congestion Impacts. *Transportation Research Record: Journal of the Transportation Research Board*, pages 22–30, 2010.

[109] I. Simaiakis and H. Balakrishnan. Dynamic control of airport departures: Algorithm development and field evaluation. In *American Control Conference*, 2012.

[110] I. Simaiakis, H. Balakrishnan, H. Khadilkar, T.G. Reynolds, R.J. Hansman, B. Reilly, and S. Urlass. Demonstration of Reduced Airport Congestion Through Pushback Rate Control. In *9th Eurocontrol/FAA ATM R&D Seminar*, 2011.

[111] I. Simaiakis, H. Balakrishnan, M. Sandberg, and R.J. Hansman. Design, Testing and Evaluation of a Pushback Rate Control Strategy. In *5th International Conference on Research in Air Transportation*, 2012.

[112] I. Simaiakis, A.D. Donaldson, and H. Balakrishnan. Impact of Heavy Aircraft Operations on Airport Capacity at Newark Liberty International Airport. In *AIAA ATIO*, 2011.

[113] I. Simaiakis and N. Pyrgiotis. An Analytical Queuing Model of Airport Departure Processes for Taxi Out Time Prediction. In *AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, September 2010.

[114] D. Simchi-Levi, J. Bramel, and X. Chen. *The logic of logistics: theory, algorithms, and applications for logistics and supply chain management.* Springer Verlag, 2005.

[115] J. W. Smeltink, M. J. Soomer, P. R. de Waal, and R. D. van der Mei. An Optimisation Model for Airport Taxi Scheduling. In *Thirtieth Conference on the Mathematics of Operations Research*, Lunteren, The Netherlands, January 2005.

[116] L. Song, H. Bateman, A. Masalonis, and C. Wanke. Integrated collaborative departure traffic management. In *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*. American Institute of Aeronautics and Astronautics, 2009.

[117] G. Soveling, J.P. Clarke, E.L. Johnson, and S. Solak. Runway Operations Optimization in the Presence of Uncertaintiest. In *Aviation Technology, Integration, and Operations Conference (ATIO)*, 2010.

[118] M.L. Spearman and M.A. Zazanis. Push and pull production systems: Issues and comparisons. *Operations research*, pages 521–532, 1992.

[119] S. Stidham and R. Weber. A survey of Markov decision models for control of networks of queues. *Queueing Systems*, 13(1):291–314, 1993.

[120] S. Stidham Jr. Analysis, design, and control of queueing systems. *Operations Research*, 50(1):197–216, 2002.

[121] M. Terrab and A.R. Odoni. Strategic flow management for air traffic control. *Operations Research*, pages 138–152, 1993.

[122] A. A. Trani, H. Baik, J. Martinez, and V. Kamat. A new paradigm to model aircraft operations at airports:the Virginia Tech Airport SIMulation Model (VTASIM). In *NEXTOR Fourth Annual Research Symposium*, FAA Headquarters Auditorium, November 14 2000.

[123] R.J. Troutbeck and W. Brilon. Unsignalized intersection theory. Technical report, Transportation Research Board, 1997.

[124] United States Government Accountability Office. Setting On-Time Performance Targets at Congested Airports Could Help Focus FAA's Actions. Technical report, May 2010.

[125] V. Vaze and C. Barnhart. An assessment of impact of demand management strategies for efficient allocation of airport capacity. *International Journal of Revenue Management*, 2011.

[126] V. Vaze and C. Barnhart. Modeling airline frequency competition for airport congestion mitigation. *Transportation Science*, May 2012.

[127] F. Wieland. The detailed policy assessment tool (DPAT). In *Proceedings of the 1997 Spring INFORMS Conference*.

[128] T.R. Willemain, H. Fand, and H. Ma. Statistical analysis of intervals between projected airport arrivals. Technical report, Rensselaer Polytechnic Institute.

[129] K. N. Yu, Y. P. Cheung, T. Cheung, and R. C. Henry. Identifying the impact of large urban airports on local air quality by nonparametric regression. *Atmospheric Environment*, 38(27):4501–4507, 2004.

[130] J. Zupan, R. Barone, and M. Lee. The Future of the New York Region's Airports. Technical report, Regional Plan Association, 2011.