# Estimating Current and Future Benefits of Airport Surface Congestion Management Techniques*

By

Alex Nakahara

B.S.E, Mechanical Engineering and Applied Mechanics

University of Pennsylvania (2010)

Submitted to the Department of Aeronautics and Astronautics on May 24, 2012,

In partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

At the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012



Author……………………………………………………………………………………………

Department of Aeronautics and Astronautics

May 24, 2012

Certified by……………………………………………………………………………………

Ian Waitz

Dean of Engineering

Jerome Hunsaker Professor of Aeronautics and Astronautics

Thesis Supervisor

Certified by……………………………………………………………………………………

Tom Reynolds

MIT Lincoln Laboratory

Thesis Supervisor

Accepted by...…………………………………………………………………………………

Eytan H. Modiano

Associate Professor of Aeronautics and Astronautics

Chair, Graduate Program Committee

* This work was sponsored by the Federal Aviation Administration under Air Force Contract No. FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

# Estimating Current and Future Benefits of Airport Surface Congestion Management Techniques

By

Alex Nakahara

Submitted to the Department of Aeronautics and Astronautics on May 24, 2012,

In partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

## ABSTRACT

Air traffic is expected to continue to grow in the future and improved methods for dealing with the increased demand on the system need to be designed and implemented. One method for reducing surface congestion at airports is surface congestion management (SCM) (also commonly called departure queue management or departure metering). The concept generally involves holding aircraft at the gate or in the ramp area instead of releasing them onto the active movement area during periods of high departure demand.

The FAA is planning to implement surface congestion management at airports where the cost/benefit analysis is favorable. Therefore, an estimate of the benefits of implementing surface congestion management in the future is necessary. To overcome the uncertainties and difficulties inherent in forecasting, this thesis adopts a multi-fidelity modeling approach and proposes three methods for estimating the benefits of SCM where the higher fidelity models study a subset of airports to inform and validate the lower fidelity models used on the entire set of airports. In the first model, a detailed analysis of a field trial of SCM at JFK airport is conducted using operational data. The second model estimates the benefits of implementing SCM at 8 major US airports from 2010 to 2030 by simulating congestion and performance levels through taxi time estimation. The last model explores several options for generalizing the results to 35 airports in the US. The results are also validated against historical benefits estimates as well as field trials of SCM where available. The findings show that SCM will result in fuel savings on the order of 1% of the total fuel burn in all stages of flight and between 5% and 45% of taxi-out fuel burn, depending on the airport.

Thesis Supervisor: Ian Waitz

Title: Dean of Engineering, Jerome Hunsaker Professor of Aeronautics and Astronautics

Thesis Supervisor: Tom Reynolds

Title: Technical Staff, MIT Lincoln Laboratory

## ACKNOWLEDGEMENTS

There are many people who have helped make my time at MIT rewarding, both academically and personally. The first is my advisor, Tom Reynolds. His patience, willingness to help, knowledge of aviation and engineering, and subtle British wit have made my work immeasurably better.

Dean Waitz and Professor Balakrishnan both provided much useful background, advice and feedback on my research. The rest of the TFDM benefits team at Lincoln Laboratory, Masha Ishutkina, Jim Kuchar, Dale Joachim, and especially Richard Jordan helped me adjust to having a desk in faraway Lexington and nudged me in the right direction many times when I was stuck.

I have to especially thank Harshad Khadilkar and Ioannis Simaiakis for helping me build on their previous work, always being available to answer my not-always-intelligent questions, and lending their considerable knowledge in the field.

My fellow labmates, especially Kristy Bishop, Chris Gilmore, Phillip Wolfe and Jamin Koo, in PARTNER are always available when I need a break, lunch, advice, or just a laugh.

I would also like to thank Tom White, Chris Maccarone and Ron Dunsky from PASSUR Aerospace for giving me access to the data from their SCM program in New York and helping with the analysis and eventual conference paper and the AeroAstro staff and faculty, especially Beth Marois and Robin Palazzolo, for all their help.

Finally, I would like to thank both my family, my girlfriend Meg Gilley, and my MIT ‘family’ for their constant support. I would not have achieved anything without you.

# Table of Contents

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ACRONYMS

| | |
|---|---|
| ADR | Airport Departure Rate (departures per hour) |
| ANOVA | Analysis of Variance |
| APU | Auxiliary Power Unit |
| ASDE-X | Airport Surface Detection Equipment, Model X |
| ASPM | Airport System Performance Metrics |
| ATC | Air Traffic Control |
| CDM | Collaborative Decision Making |
| CDQM | Collaborative Departure Queue Management |
| $CO_2$ | Carbon Dioxide |
| DMAN | Advanced Departure Manager |
| FAA | Federal Aviation Administration |
| FACT | Future Airport Capacity Task |
| FCFS | First-Come, First-Served |
| ICAO | International Civil Aviation Organization |
| IMC | Instrument Metereological Conditions |
| NAS | National Airspace System |
| OEP 35 | Operational Evolution Partnership (top 35 commercial airports with significant activity) |
| OOOI | Out-Off-On-In times (gate OUT, wheels OFF (take-off), wheels ON (landing), gate IN) |
| RF | Random Forest |
| SARDA | Spot and Runway Departure Advisor |
| SCM | Surface Congestion Management |
| SWAC | System-Wide Analysis Capability |
| VMC | Visual Metereological Conditions |

# 1. INTRODUCTION

Air traffic is expected to continue to grow in the future and improved methods for dealing with the increased demand on the system need to be designed and implemented in all phases of flight. The airport surface is one area where system inefficiencies are especially evident in the form of congestion: at the OEP 35 in the United States in 2010 there were over 48 million minutes of departure taxi delay [1] (i.e., time over the unimpeded time), translating to over 194 million gallons of excess fuel burn. One approach to mitigate the resulting monetary and environmental impacts is to employ surface congestion management techniques (also known as departure queue management or departure metering). Understanding the potential benefits of these techniques is important to help prioritize them relative to other capabilities which could be developed to help address future air transportation system needs. However, these benefits are difficult to calculate because the performance (and therefore congestion) of an airport is dependent on a variety of factors such as capacity, weather, demand, configuration, and controller performance, among others. The work described in this thesis develops methodologies and applies them across a range of study airports to assess the potential benefits of surface congestion management techniques under current and future operations to help inform decision-making for future air transportation system evolution.

## 1.1 SURFACE CONGESTION AND ITS IMPACTS

Every airport can be considered to have a limit to the number of aircraft it can efficiently handle in a given time period as a function of characteristics such as configuration, weather conditions and demand. When demand increases above the level of the airport's capacity at a given time, congestion starts to grow. There are methods for managing congestion, such as the slot-control system in place in Europe [2]. This system allocates a certain number of departure slots per hour to airlines and forbids the scheduling of additional flights beyond the amount of slots. The number of slots is tied to the bad-weather IMC capacity of the airport. While congestion cannot be entirely eliminated, this system does substantially reduce it. However, the US has no such system (excepting the New York area airports and Washington DC Reagan at some times), meaning that schedules are often created by airlines assuming best-case VMC, resulting in significant congestion when capacity is lower (such as during bad weather), which must then be managed in "real time". The FAA projection of continued growth in demand would result in unsustainable levels of delay and congestion without improvements to the existing system. Surface congestion negatively impacts airports in several ways. It is a major source of delay, which can propagate from one airport through the entire system. It is environmentally and fiscally wasteful, causing excess fuel burn, pollution, and delay costs to airlines, passengers and local/national communities impacted by the airport's activities. It also results in an increased workload on air traffic controllers because there are more flights actively taxiing. While eliminating all of these impacts is not realistic, SCM is designed to mitigate many of them.

Figure 1 shows a particular airport with historically high delays, John F. Kennedy Airport in New York, and how both the average taxi time and number of flights with taxi times greater than 40 minutes are strongly correlated with demand levels. In addition, it shows how demand consistently rose from 2000 until 2008, the start of the global recession. The forecast demand until 2030 is also shown, and the peak level of 2007 is quickly passed.

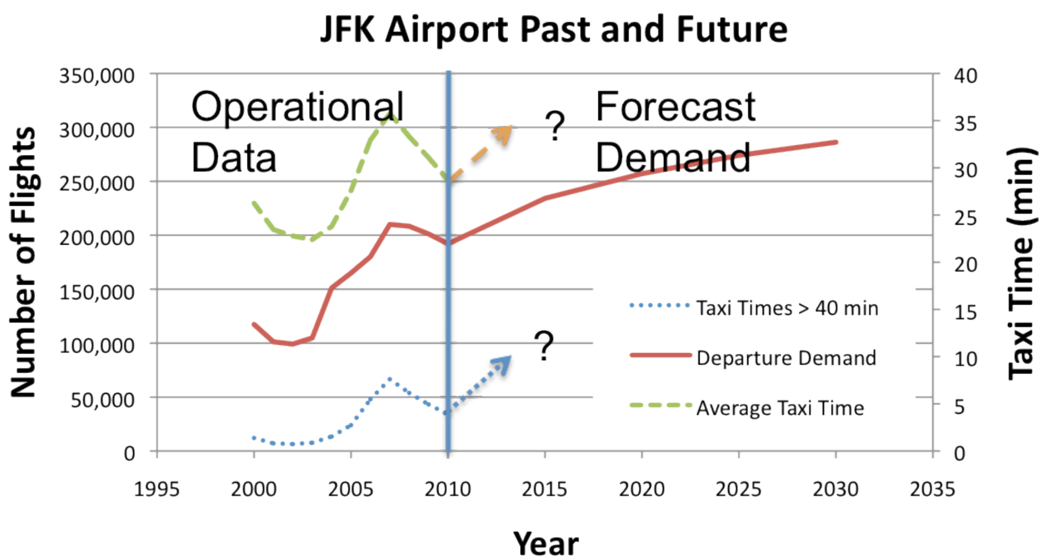



**Figure 1: JFK Behavior**

Figure 2 shows that these trends are not limited to a few highly congested airports but are present across the OEP 35 airports. By 2030 the FAA projects a increase in air traffic of 80% over 2010 levels at the OEP 35 airports [3]. While a much smaller proportion of flights are operating in congestion (which can be viewed as the number of flights with excessive taxi times, and multiplied by 10 here for better visibility) than at JFK, congestion grows quickly with a relatively small increase in demand because many of the major US airports are already operating close to their capacity.

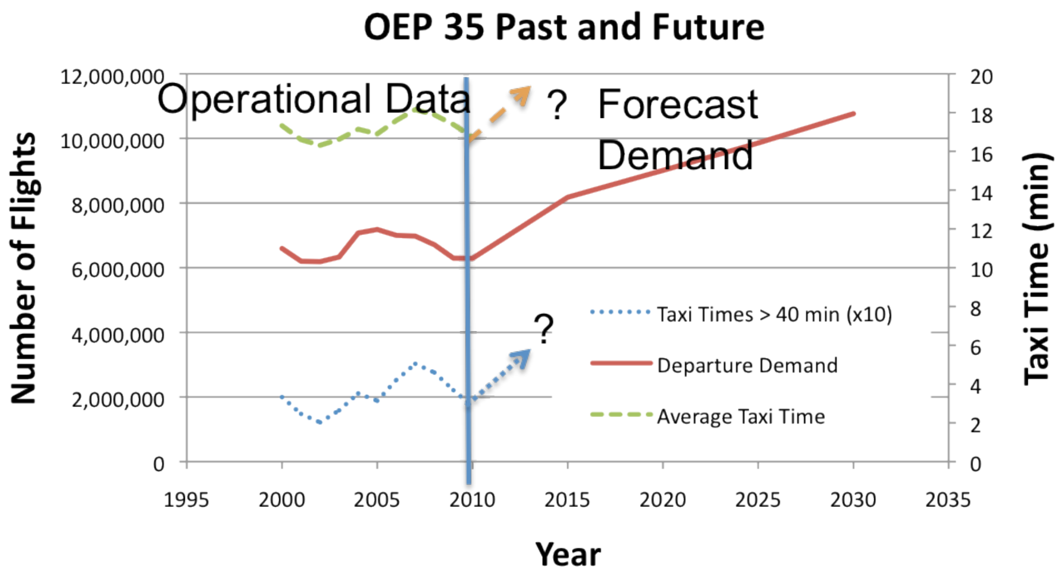



**Figure 2: Relationship between Demand and Taxi Time for OEP 35 Airports**

The relationship between demand, capacity and congestion is shown in Figure 3. The figure demonstrates how delay varies with ρ, which is the ratio between demand and capacity. As ρ approaches 1 (for an extended period of time), delay increases nonlinearly [2] so that an incremental increase in demand results in a large increase in delay. This relationship will become more and more relevant given the constant increase in demand shown in Figure 2 in the future. The forecast demand is larger than any seen in the last 10 years and indicates that congestion is a problem that will need to be addressed at a system level. One of the questions that this thesis attempts to answer is what that increase in demand will mean for taxi times and levels of congestion with and without future mitigations such as surface congestion management.

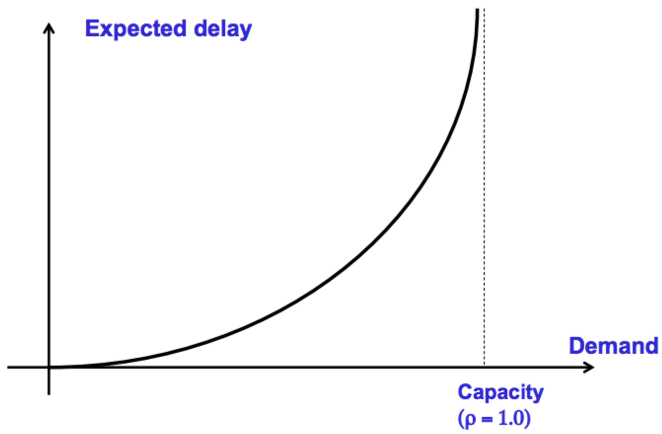



**Figure 3: Relationship between Demand, Capacity and Delay[3]**

## 1.2 SURFACE CONGESTION MANAGEMENT

Surface congestion management is a tool that can reduce some of the impacts of congestion such as fuel burn and emissions by reducing the time flights spend taxiing with their engines on. Additionally, although it reduces congestion it is designed to keep the airport operating at its maximum capacity during periods of high demand. In generic terms, SCM achieves this by identifying an efficiency threshold in terms of number of flights that the airport can efficiently handle at one time. If the airport is below this threshold, no additional management is needed. When the number of flights seeking to depart exceeds this threshold, excess flights are not allowed to push back from the gate and instead are held at the gate or some other appropriate location with engines off until they can be released to the departure runway more efficiently, as shown in Figure 4. By restricting the number of active flights in this way, "engines-on" taxi-out time, fuel burn and emissions can be reduced (as long as other operational requirements are still met which can vary by airport).

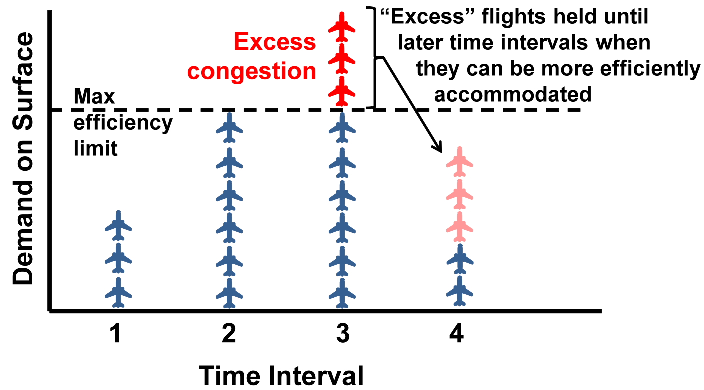



**Figure 4: Surface Congestion Management Concept**

Note that this is largely simply shifting the delay from being incurred on the taxiways with engines running, back to the gate (or other designated hold location) where the engines can be turned off. However, in addition, it is possible that SCM could lead to some net delay reduction as well at severely congested airports because of the reduction in controller and pilot workload as well as the non-linear relationships in congestion between taxi time, throughput and traffic levels. While these effects are not studied in this thesis they would be a good subject for future work.

A basic concept used in this thesis to quantify congestion will now be introduced. A useful way of visualizing the performance of the airport as a function of the surface traffic is the *throughput saturation curve,* developed by Shumsky [10] and Pujet [11], and illustrated in Figure 5. It represents the departure rate as a function of an appropriate surface metric (such as the number of aircraft taxiing out or in the departure queue). Different traffic metrics might be appropriate for different airports, but typically for low levels of surface traffic, as more departing aircraft are pushed back onto the surface, the departure rate increases as more aircraft are available at the runways.  However, as surface traffic increases further, the capacity limit of the airport is approached and the departure rate eventually saturates. The saturation throughput is not the maximum throughput observed because saturation curves are often not well-behaved and to increase the throughput from the saturation throughput to the maximum usually requires a disproportionate increase in congestion.

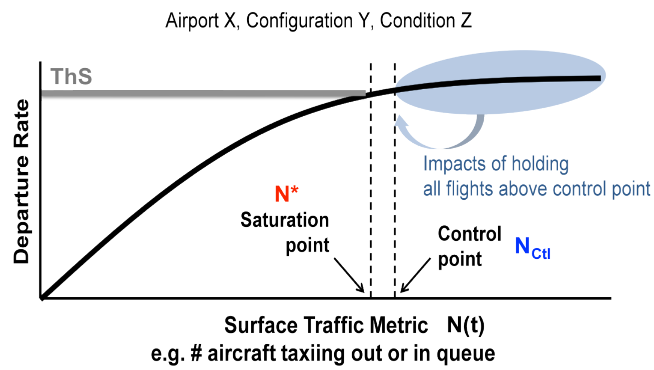



**Figure 5: Airport Saturation Curve**

At the saturation point, any additional increase in surface traffic simply adds to congestion and does not achieve any substantial increase in the departure rate (indeed, if surface traffic gets very high, the departure rate can decrease due to gridlock). Archived operational data can be used to determine saturation curves for different airports under different configurations, weather and traffic characteristics. These curves can then be used as a basis for when to hold aircraft from taxiing when the airport is expected to operate above some control point on the curve. Typically the control point would be slightly higher than the expected saturation point so as to avoid risking loss in departure rate, but not so large as to lose significant benefits from the control strategy. The impacts of SCM can then be assessed in terms of the performance implications (e.g., taxi time and fuel burn) from moving the operating point of the airport from above the control point back to the point, as shown in Figure 5. Saturation curves will be used for many of the analyses in this thesis and some key terms will be defined here (shown in corresponding colors in Figure 5):

Number on Surface (N(t)) – Traffic Metric

N* - Saturation point

$N_{Ctl}$ – Control Point

Saturation Throughput (ThS) –Departure rate at N*

## 1.3 TYPES OF SURFACE CONGESTION MANAGEMENT

Although the principles of surface congestion management as described above are generally applicable across congested airports, the specifics of implementation at any given airport (e.g., how to determine when flights need to be held, coordinate which specific flights to then hold, where to hold them and what level of airline interaction is needed) depends on the airport/traffic characteristics and the level of sophistication desired. To illustrate this point, several specific implementations of surface congestion management have recently been tested in field trials or in simulation environments, results from which will be described in detail in Chapter 2.

Pushback rate control (or N-Control) has been tested at Boston Logan Airport which recommends a general pushback rate to controllers to limit the number of aircraft on the surface at peak times [4]. The pushback rate is explicitly informed by N(t); hence the alternative name of N-Control. Specifically, N(t) is monitored in real time (and projected into the near future) and pushbacks are suspended if N(t) rises or is expected to rise over a threshold value. The method was adapted slightly to account for controller preferences by assigning suggested pushback rates in 15 minute intervals.

CDQM, tested at Memphis airport, allocates departure slots to different airlines at peak times to manage average departure queue delay to below a control value, and then the airlines determine which flights go into which allocated slot [5]. Airlines are allocated slots according to ration-by-schedule, which allows flexibility and prioritization of flights.

Another class of approach (subsequently referred to as the PASSUR method, because they have implemented such an approach at JFK airport) recommends when specific flights should leave from gate or spot to manage surface congestion [6]: this affords greatest control (and hence potentially the greatest benefits) but requires significant real-time airline coordination to know when flights want to push-back, as well as communication of, and compliance to, allocated slot times which may be later than the desired push time in order to better manage the demand when it exceeds the capacity of the airport.

These three approaches have an increasing level of complexity and correspondingly have an increasing prediction horizon, with pushback rate control issuing advisories for only the next time period while the PASSUR method sets initial pushback times a day in advance (although these times are modified through the day as circumstances dictate). Despite these differences they can all be abstracted as a form of throughput saturation curve because they all seek to restrict the amount of planes actively taxiing (or in physical queues) to a control value, implicitly or explicitly. While CDQM explicitly controls the delay time, this can be related to first order to the length of the departure queue. Metering individual flights is more complicated, but essentially restricts the size of the queue to a smaller amount. At its root, any implementation of

SCM aims to reduce taxi time as much as possible while keeping throughput at its peak value. Saturation curves are a way of determining the maximum number of flights that can be held off the surface before throughput is substantially affected, which is the same result because taxi time can be related to the number of actively taxiing flights.

*While different airports might require different specific implementations of SCM, we assert that saturation curves can provide a first-order benefits estimate for any airport because all types of metering can be idealized as variations on N-control.*

To provide validation for this assertion, the results from field trials of each of the approaches will be examined and compared to the results obtained from the saturation curve method that is being used in this work.

## 1.4 NEED FOR BENEFITS ASSESSMENTS

The reason that we are calculating these benefits (and characterizing different types of metering as N-Control) is that there is a strong need to identify the most cost-effective options for dealing with increased demand in the future. There are many possible improvements such as SCM to current air traffic management technologies and techniques that are required to move toward the next generation air transportation system (NextGen). Other upgrades include controllers being supplied with advanced surveillance and flight data management display systems that will allow them to maintain an integrated picture of the current situation. Controllers and supervisors may also be provided with a suite of Decision Support Tools (of which an SCM tool could be one element) that provide critical information for assistance in tactical and strategic decision-making. In addition, NextGen capabilities will facilitate data exchange between controllers within a tower facility, between ATC facilities, and between stakeholders such as airlines.

The capabilities provided by these systems should enable multiple system benefits, such as reduced surface delay, taxi time and fuel burn (with associated improved operational and environmental performance); better performance during severe weather and other off-nominal conditions; improved usability and situational awareness; and enhanced safety. However, in order to assess the viability of specific tools for NAS-wide deployment it is necessary to undertake a cost-benefit analysis. This includes estimating the likely costs of deployment at appropriate locations relative to the potential benefits this deployment will bring over several decades of operation. The generation of such data is also highly complementary to the prototyping effort that is often conducted as part of advanced system deployment. For example, the process by which benefits are identified necessarily requires an understanding of the inefficiencies present in the current baseline ATC system. Understanding the causality of these inefficiencies can help identify what capabilities are needed to address them, and therefore, help guide priorities for the prototype system.

Estimating the future benefits of any system is at best an uncertain task. SCM presents several challenges in particular: identifying a methodology that is robust enough to accommodate the many changes in airport behavior and still be valid into the future (saturation curves), determining how saturation curves change over time because they can vary based on demand levels and airport usage, and accurately predicting the level of congestion (which is inherently uncertain) are all major issues that had to be overcome. In addition, the forecasts upon which the assessment is based have their own uncertainty.

## 1.5 SUMMARY AND ORGANIZATION OF THE THESIS

In order to better understand the role surface congestion management can play in the air transportation system and to make the case for its deployment at different airports, a benefits assessment is required. Current benefits assessments are restricted to the present-day, so a new methodology for future benefits assessments is needed to assist with decision-making regarding what systems to develop and deploy in the air transportation system over the next several decades. *Recognizing the need for a system-wide benefits estimate of SCM and that airports have unique and distinct operating characteristics that make it difficult to develop a generalizable method, this work adopts a multi-fidelity modeling approach where the higher fidelity models at a subset of airports will inform and validate the lower fidelity models which are then applied more generally.* This approach is depicted in Figure 6.

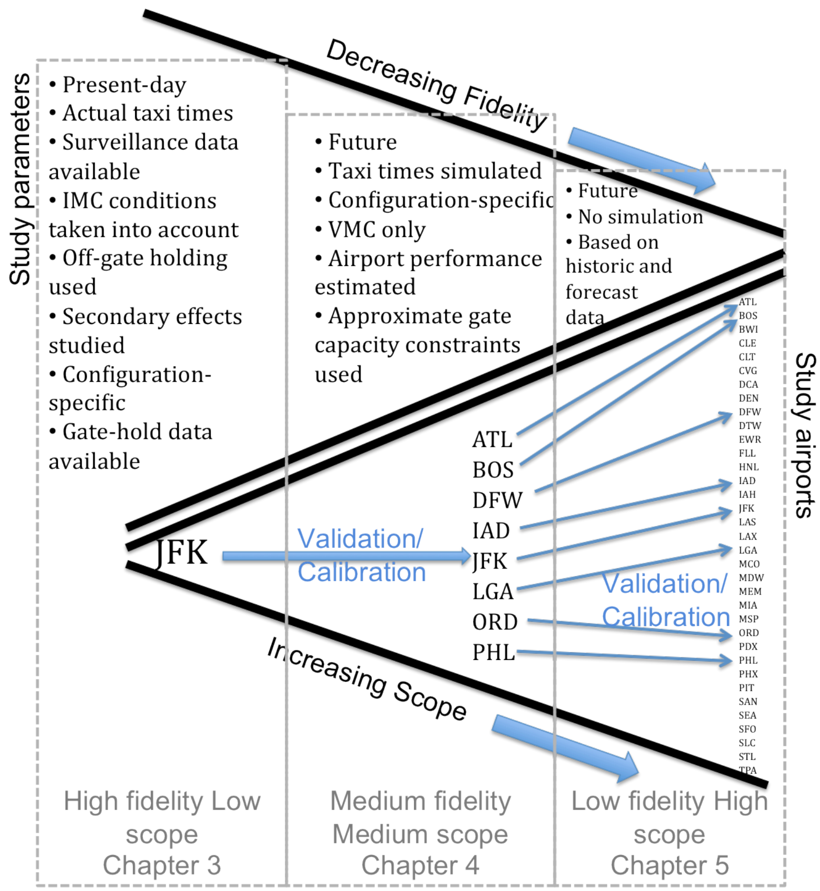



**Figure 6: Multi-Fidelity Modeling Approach**

This thesis will:

- Examine previous benefits assessment methodologies and field trials for current day operations in Chapter 2.
- Present a high fidelity analysis for estimating the benefits of SCM from a 6 month field trial at JFK airport in Chapter 3 for comparison to other field trials and to aid in construction of medium and low fidelity models.
- In Chapter 4, present a medium fidelity methodology for estimating future performance (as measured by throughput) and congestion at several key airports and derive benefits estimates. Additionally, examine impact of key operational constraints (in the form of gate availability) on the benefits from SCM and validate results by comparing to field trials and the high-fidelity method.

- In Chapter 5, present several low fidelity methodologies for estimating the benefits of SCM across the NAS in the future, validated by the results from the 8 medium-fidelity study airports. Compare and contrast methodologies and benefits estimates and draw insights on key variables that identify airports that have large possible benefits from SCM.
- In Chapter 6, review key results and discuss the implications for policy makers.

# 2. PREVIOUS WORK

## 2.1 BACKGROUND

Benefits analyses in the literature are mainly focused on present-day operations which shed little light in terms of benefits assessment for future system deployments: benefits assessments under future operations are critical to this process. Several methodologies have been proposed for calculating the effect of metering on current-day airports using operational data such as ASPM or ASDE-X. In addition, three field trials have been performed in Boston, New York, and Memphis using separate methodologies. This chapter examines the different methodologies and assesses their applicability for estimating future benefits. In addition, the benefits shown in the field trials are discussed and compared to a theoretical approach using saturation curves.

## 2.2 THEORETICAL MODELS AND SIMULATIONS

Several distinct methods for calculating the benefits of SCM have appeared in the literature. The first is a method that will be used in this paper, developed by Simaiakis [9], Pujet [10] and Shumsky [11]. Linking the number of aircraft on the surface $N(t)$ with the departure throughput of the same time period results in a relationship that predicts when an airport will be in congestion. The benefits of performing N-control metering can be calculated by comparing the taxi times in congestion and at saturation as was explained previously. We claim that this method can also be used to give a first-order estimate of any type of metering.

CDQM allows for a simple analysis of benefits because it is controlling to a target amount of queue delay. In [5], the benefits of using CDQM at Memphis for an entire year are calculated by assuming that taxi times greater than the sum of the average unimpeded time plus the target delay are excess taxi times that would be reduced through CDQM. While this is simple on the surface to calculate, there are issues that make it more complex when expanded to other airports, such as the proper unimpeded time to use (can vary with configuration and gate location) and the target delay (unique to an airport / configuration and tailored to ensure maximum throughput). In addition, if the delay is high enough (or if the airport geometry requires runway crossings or queues besides the departure queue) the assumption of unimpeded taxi to the departure queue might be not accurate. Finally, the ASPM database does not provide a breakdown of taxi time, making it difficult to determine the cause of high taxi times. Two examples of high taxi times not caused by congestion are de-icing airplanes during the winter and flights receiving ground holds due to weather at their destination. Using the saturation curve approach reduces the impact of such flights because it looks at average taxi times at the saturation point and in congestion.

Because operations like de-icing take place during both low and high periods of demand, the impact on the average taxi time is spread equally to the congestion average and the saturation average. The benefits then do not depend on such occurences.

The CDQM approach of setting a target delay can also be thought of as a derivative of the N-control method. The target delay should ideally be the minimum time above the unimpeded time that is necessary to maintain maximum throughput. As was mentioned previously, some queues (and therefore taxi times greater than unimpeded times) are necessary to avoid the runway being starved. But this target delay (plus the unimpeded time) should be the taxi time at N*. Given this, the approach becomes identical to the saturation curve method. However, calculating the optimal target delay is not as straightforward as calculating N*.

The third method was developed primarily by researchers from Sensis Corporation (with input from the FAA) [13,14,15] as a way to measure the theoretical benefits from implementing a departure management tool (DMAN). [12] was an analysis of ASDE-X surveillance data from 2008-9 at JFK airport. By assuming that the average time spent per unit length in the 'departure queue' (defined as the physical queue for the departure runway) does not vary with the length of the queue, the benefits of restricting the queue to a given size can be calculated by subtracting the idealized taxi time (unimpeded + time spent in a short queue) from the total taxi time. This assumes that aside from the time spent in the departure queue there is no other source of taxi delay. The benefits were calculated for departure queues controlled to lengths of 5 and 10, corresponding to aggressive and conservative values.

In [13] a taxi simulation was developed based on ASDE-X data that allowed for a version of DMAN to be directly tested. Two scenarios were simulated, one with DMAN implemented and the other for operations as currently run (FCFS). Because the simulation allows all other factors to be kept the same, the differences in taxi times between the two scenarios are the benefits from implementing DMAN. Their proposed implementation (DMAN) of SCM includes resequencing departing flights to improve throughput so the benefits are not quite comparable to the method of Simaiakis. In addition, their methodology is based on limiting the physical departure queue at the runway instead of the overall number on surface. They used ASDE-X surveillance data to calibrate their simulation, which models the motion of each flight in the movement area with interactions with other flights, as well as forming queues if necessary. This method has the benefit of transparency and easy interpretability: because the same day is being simulated with and without SCM, one can be surer of the effects of SCM. The difficulty in extending this method to the future is the lack of availability of ASDE-X data at all airports, the presumed customization necessary to run the simulation at multiple airports, and possibly time constraints. In [13] Stroiney et al. examined one day at PHL and JFK; to obtain a more robust estimate across different configurations and weather conditions a larger sample size would be desired. In [14] Stroiney and Levy extend their analysis of the benefits to multiple airports as well as account for gate constraints by examining the same day as [13] at JFK. Their conclusion at JFK was that limits on gate space result in a small and ultimately negligible decrease in benefits from SCM. To extend their analysis to multiple airports, they observed that, akin to the saturation curve method, SCM benefits are roughly equal to the amount of time spent by flights taxiing when the departure queue is above its target length. The benefits are then obtainable by observing the departure queue length over a time period. This method is only feasible with recorded surveillance data, making future prediction impossible.

To add to these methods, we will propose a new method in Chapter 3 that is based on observations of actual metering operations. ASPM and ASDE-X taxi time data is compared to the schedule of gate holds performed to obtain the benefits of the field trial conducted by PASSUR at JFK in 2010. The first method will also be used as a validation in Chapter 3, as well as the main element of the analysis in Chapter 4.

Although beyond the scope of the current study, there are also longer-term NextGen tools under development that go a step further and aim to reduce overall delay by combining SCM with wider airport surface optimization decisions (such as configuration selection, aircraft sequencing and taxi routing). For example, in terms of sequence optimization, by ordering flights into an optimized queue based on the size of the aircraft, throughput can be increased due to the different separation standards between differently sized aircraft. These advanced concepts assume that taxi times can be accurately estimated. With a predetermined sequence and knowledge of the taxi time of a given flight, pushback times can be given to every flight so that it can travel unimpeded to the runway and immediately take off. Examples of this include CDM in Europe [7] and SARDA (NASA) [8].

## 2.3 FIELD TRIAL STUDIES

We will now examine the results from several field trials in detail both to gain insight on the mechanics behind SCM as well as to use for validation of our analyses. Because the results from BOS are presented for portions of specific days, we also develop an extrapolation technique that yields an annual estimate of benefits from SCM for better comparison with other methods.

### 2.3.1 Boston Logan (BOS): N-Control

**Background**

Simaiakis et al. [4] demonstrated a different method of SCM at Boston Logan in 2010, *Pushback Rate Control* (also referred to as N-Control). Pushback rate control follows a simple heuristic: if the total number of aircraft taxiing to a departure runway ($N$) exceeds a control value ($N_{Ctl}$), further pushbacks are stopped until $N$ is below the threshold. This method is informed by the saturation curves mentioned earlier in this paper.

The specifics of the strategy used by Simaiakis et al.were developed through discussions with the BOS facility as well as study of historical data from the airport. Pushback rates were suggested to controllers in 15 minute intervals based on the current level of congestion on the surface, allowing some variation between no restrictions and a full stop on pushbacks. Due to the complex nature of the runway layout at BOS (see Figure 7), there are many different configurations in use. Because the parameters (such as N* and throughput) that affect the SCM strategy vary depending on the configuration, only the 3 most common configurations (accounting for 70% of use) were used. In addition, IMC conditions were not considered because of separate procedures that are followed at BOS during such periods.

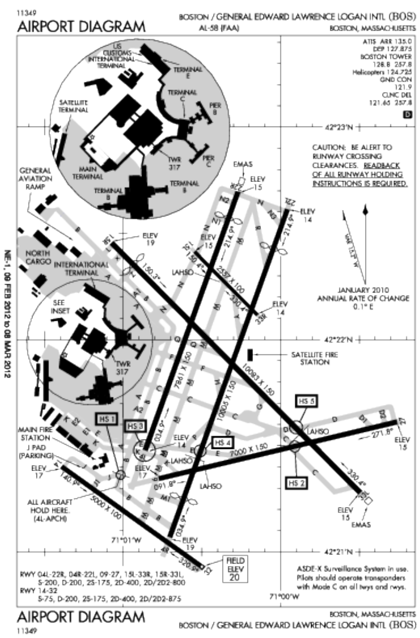



**Figure 7: BOS airport diagram, showing alignment of runways [25]**

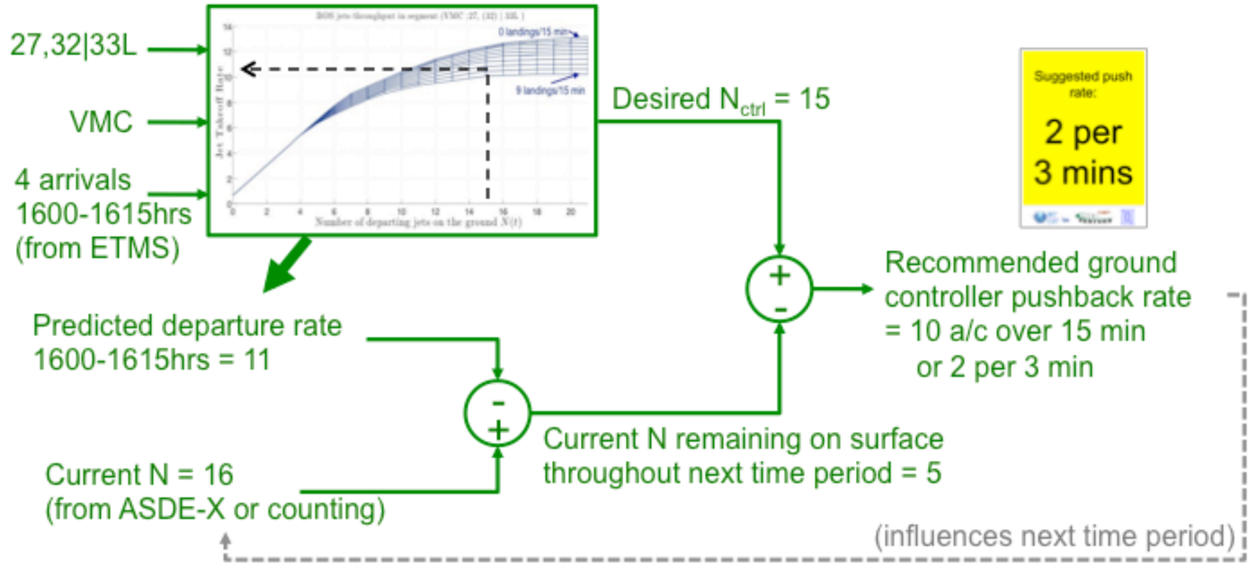



**Figure 8: Schematic of pushback rate calculation [4]**

A diagram showing the determination of the pushback rate is in Figure 8. The saturation curve (top) is used to determine both the desired control value as well as the estimated throughput for the next period. The estimated throughput is subtracted from the current N value to estimate N after the next period. This can be compared to the control value to derive the recommended pushback rate for the current period.

**Results**

The field trial was conducted over a month from August 23 to September 24, 2010 on select days during the evening departure push period (4-8PM). There were a total of over 37 hours where metering was in effect, and 24 hours of test periods with significant gate holds. Furthermore, it was found that one configuration experienced almost no metering (4L, 4R | 4L, 4R, 9) with the other two main configurations (27, 22L | 22R and 27, 32 | 33L) receiving the bulk of the congestion.

An important assumption made in the study was that each minute of gate hold represented a minute of saved active taxi time. While this was shown in section 3.2.4 to not necessarily be true at JFK, it is valid here because of the short duration of the holds performed at BOS. Simaiakis et al. [4] reported a total of 247 held flights over the 8 days with significant amounts of gate holds at an average hold time of 4.35 minutes, yielding a total of 1075 minutes of gate hold (and saved taxi time).

### Extrapolation to Annual Benefits

The results from the BOS field trial are not easily comparable to the results from other field trials as well as the results expected from analysis of ASPM data because metering was not in effect every day in the month or even for the entire day on the selected days. In order to use the BOS field trial as validation for this thesis, an extrapolation to yearly benefits was needed. Therefore, we extrapolate the results presented by Simaiakis et al. [4] to provide an estimate for implementing N-Control for an entire year at BOS.

There are many possible ways to extrapolate, including scaling by time (24.5 hours of metering compared to 4*365 hours of peak evening traffic), scaling by demand, scaling by time in a specific configuration, etc. We chose to decouple the time window chosen (4-8 PM) from the assumed high demand levels during that time. We defined a "high demand period" as a period in which there were more than $X_i$ departures in a fifteen minute period, where $X_i$ is the fewest departures in a period witnessed during the time when metering was in effect and $i$ represents the configuration in use. For 22L, 27 | 22R $X_i$ was found to be 10 departures / 15 minutes and for 27, 32 | 33L it was found to be 8 departures / 15 minutes. We then found all other periods when the airport had more than $X_i$ departures **and** was in the same configuration with VMC conditions. Because 4L, 4R | 4L, 4R, 9 required very little metering, only the two other configurations were examined. Benefits are also dependent on the level of demand so the extrapolated benefits were scaled by the ratio between the annual number of flights in high demand periods and the number seen in the trial. For 22L, 27 | 22R there were 525 departures during the trial and 8,645 over the course of 2010, resulting in a scaling factor of 16.47. Note that this is larger than might be expected if the benefits were simply scaled by time (one month of metering * 12 months in the year). This is due to time periods outside of the 4-8 PM study period experiencing congestion, as well as the variation of airport behavior and configuration choice over the year. The scaling factor for the other configuration is even higher, at 23.69. These scaling factors result in annual benefits estimates for the two main configurations of 150 hours for 27, 32 | 33L and 191 hours for 27, 22L | 22R. These are much lower than the configuration-specific benefits seen in the JFK field trial where the benefits were in the thousands of hours, but are to be expected given the difference in the prior demand and congestion levels at the two airports.

These totals can be validated against the estimates obtained from the saturation curve method for these two configurations. $N_{Ctl}$ is assumed to be equal to $N^*=17$ for both configurations. This value is different from the N* given by Simaiakis because it is calculated using the ASPM data instead of the ASDE-X data he used. ASDE-X records aircraft from the spot to the runway while ASPM records aircraft from the gate to the runway, resulting in a higher N* for ASPM. Figure 9 shows the benefits from the saturation curve method in blue compared to the extrapolated field trial benefits in red. The variance in the blue bars represents the difference in increasing or decreasing $N_{Ctl}$ by one aircraft. Figure 10 shows the benefits from the saturation curve method for the 10 most used configurations in VMC at BOS in 2010 as well as the corresponding $N^* = N_{Ctl}$ value. Note that the estimated benefits for metering in the 4L, 4R | 4L, 4R, 9 configuration is only 20 hours which agrees with the observation Simaiakis made that that configuration rarely experiences congestion.

Despite the relatively short duration of the BOS SCM trial, it still serves as a useful validation of the saturation curve method for estimating the benefits of SCM. In addition, it agrees with several observations made during the JFK study such as that benefits are strongly tied to the configuration of the airport.

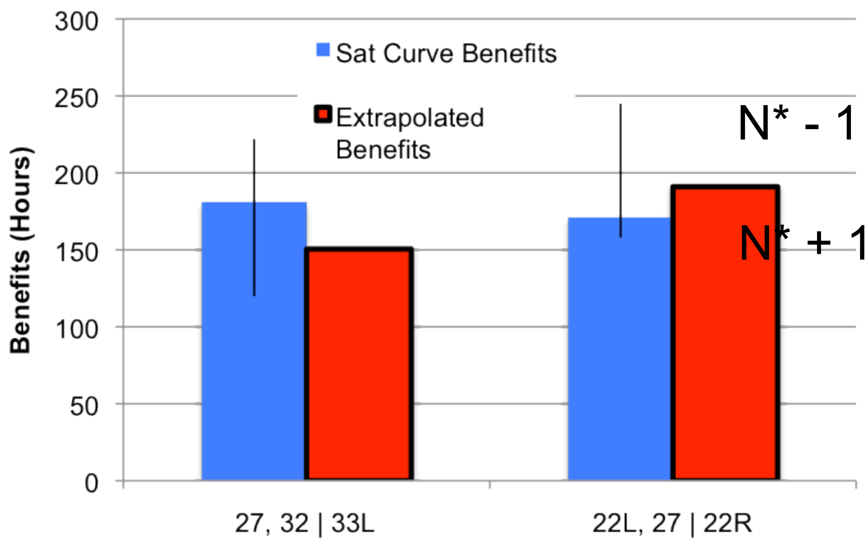



**Figure 9: Comparison of Benefits Estimates**

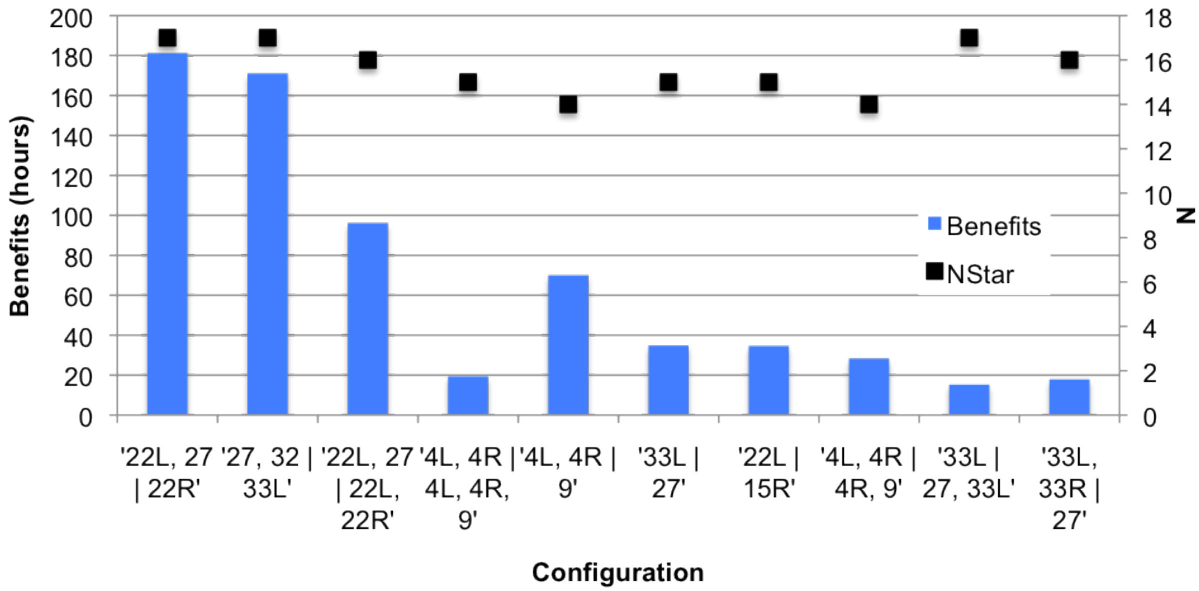



**Figure 10: Saturation Curve Benefits across major configurations**

### 2.3.2 Memphis (MEM): CDQM

CDQM field trials were held on 20 days in 2010 at MEM. For half of this period, Delta was the only participant, with FedEx joining in August. [5] examines in detail 2 days as case studies and notes that when the airlines follow the recommended slot allocations, the delays are generally at the level of the target delay of 6 minutes (as opposed to a day where the slot allocations were calculated but not used, when the delays reached 20 minutes).

Quantitative benefits for the Memphis field trial are not given in [5], but the time spent in queue is shown to be qualitatively held to the target delay of 6 minutes. Based on this [5] assumes that CDQM would be successful in reducing taxi times to at most the unimpeded time plus the threshold delay time. By capping actual taxi times in 2008 to this value, they claim annual benefits of 86,000 minutes (1,433 hours). As a comparison, we can calculate the benefits of SCM using the saturation curve method (Saturation curve used is shown in Figure 11). Using an $N_{Ctl} = N^* = 25$ produces benefits of 87,200 minutes (1,454 hours) for 2008. This similarity demonstrates the interchangeability between the saturation curve method and target delay method for calculating the benefits of SCM.

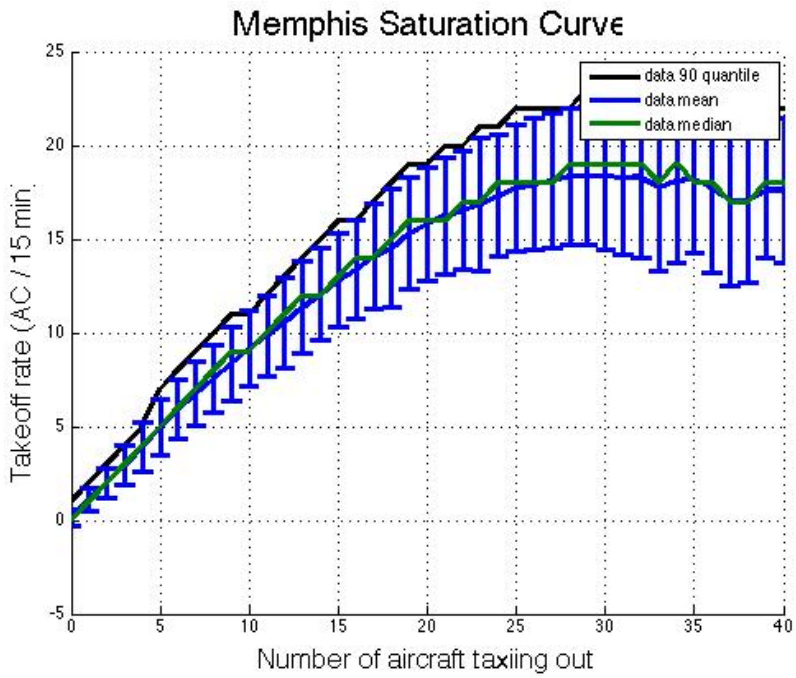



**Figure 11: Memphis Saturation Curve – 2008**

### 2.3.3 New York Kennedy (JFK): PASSUR

A basic type of SCM has been in place at JFK for use during severe weather conditions, and has recently been expanded to be used constantly. It is not explicitly a saturation curve-based approach, as it assigns hold times to individual flights. There were no existing benefits assessments that could be used to compare with the saturation curve method. In addition, the large amount of data available (6 months) provided the opportunity for a high-fidelity analysis to examine the primary and secondary effects of SCM. For these reasons, the analysis presented in Chapter 3 was performed.

## 2.4 CONCLUSIONS

Several methods for calculating benefits from SCM in literature were reviewed. The methods proposed by Sensis Corporation [13, 14, 15] use operational ASDE-X surveillance data and are therefore inappropriate for use in the future. CDQM [5] was shown to be similar to the final method developed by Simaiakis [9] but with the added step of calculating the target delay, making the Simaiakis method of saturation curves the most appropriate to adapt for benefits assessments looking at the future.

In addition, results from field trials at BOS and MEM (JFK will be examined in detail in the next chapter) were discussed and compared to the estimates derived from saturation curves. The results were comparable, supporting our assertion that the saturation curve method can be used to approximate different types of SCM.

# 3. HIGH FIDELITY ANALYSIS FOR ASSESSING SCM BENEFITS AT ONE AIRPORT

The methods listed in Chapter 2 approach the estimation of benefits from SCM from several different directions. However, all are theoretical and should not be relied on without detailed validation from real-world data. While the field trials at BOS and MEM can be used as validation, the relative brevity of the trials is not ideal. In addition, an approach is needed that represents a third form of SCM that is not explicitly (BOS) or implicitly (MEM) based on saturation curves to support our assertion that saturation curves can be used to model all types of SCM. To provide a more in-depth validation, we calculate and examine in detail the results from a field trial of SCM at JFK airport in New York City. The methodology used is necessarily different from that in the BOS field trial, where it was assumed that 1 minute of hold time resulted in 1 minute of saved taxi time. While this assumption is valid at BOS because of the short hold times, it could not be assumed at JFK because the holds were longer and were often conducted off-gate. This introduced small (but measurable) losses when compared to the 1 to 1 metric.

The JFK analysis also acts as a high fidelity model to validate and support development of models at medium and lower fidelities with wider applicability. Several secondary effects of metering such as throughput and physical constraints are examined in detail.

## 3.1 BACKGROUND

JFK is one of the biggest and most congested airports in the US, with the highest average taxi time in the nation in 2009 (31 minutes) [1]. The layout of the airport is shown in Figure 12. Early forms of surface congestion management have been used at the airport since 2002 to assist with deicing operations. In February 2010, a full-time implementation of prototype software and processes was put in place by PASSUR Aerospace for the Port Authority of New York and New Jersey, initially to manage the disruption caused by a five month closure of one of the major runways (13R/31L) at the airport. However, its use was continued when the runway re-opened.

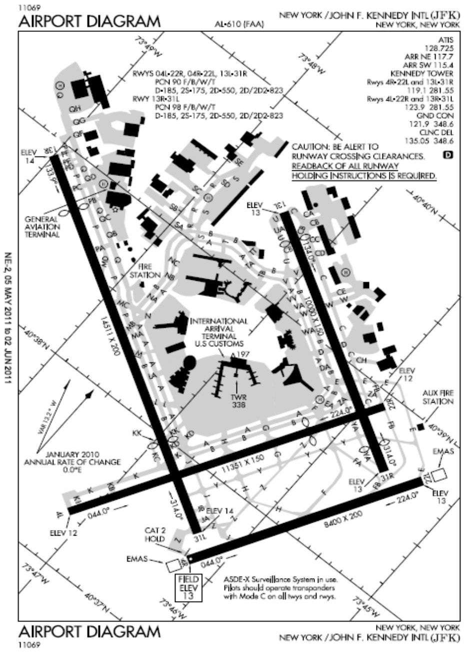



**Figure 12: New York JFK Airport [25]**

A schematic of the implementation of the Surface Congestion Management (SCM) approach at JFK is shown in Figure 13. The development of the approach was based upon a collaborative process in which all carriers participated to ensure the maximum use of departure capacity while reducing the amount of engines-on departure taxi time. One of the cornerstones of the approach was the use of predictive analytics to accurately forecast up to eight hours in advance the expected departure and arrival capacity (in terms of departure and arrival "slot counts") of the airport based on the weather forecast and past airport performance under identical predicted weather conditions. This in turn was used with the demand information of flight-specific requested push-times sourced from (and updated by) the airlines to develop the initial allocation of flights to permitted taxi "slot times" over the forecast period. When the number of aircraft wanting to push-back was below what the airport could efficiently handle in a certain time period, the slot times were the same as the desired push times. But when the number of flights wanting to push exceeded what the airport could efficiently handle, the excess flights were allocated slot times later than their desired push times to better manage the demand. The initial allocation of flights to slot times used the concept of "ration by schedule" [15] in which the number of slots per hour was allocated to each operator based on their normal (unrestricted) percentage of the hourly volume. Slots were issued up to two hours in advance, to accommodate the longer planning horizon of international operations. Once the initial allocation of departure slots had occurred, the users had the opportunity to request swaps and substitutions within their allotment of departure slots, in order to better reflect their internal business priorities. These requests were received and processed electronically via a web interface managed by the "slot allocation manager": a neutral third-party established to run the program. All slot assignments could be seen by all program users, ensuring maximum transparency and trust that there was no gaming of the system. The central tenet of the above process is that users do not push-back until they have reached their assigned departure slot time rather than simply pushing back whenever they are ready (i.e., as happens when surface congestion management is not in effect). When a flight's slot time was later than the requested push time, the hold time was absorbed either at the gate or, if the gate was required by another aircraft, at a pre-assigned holding pad with engines off as much as possible.

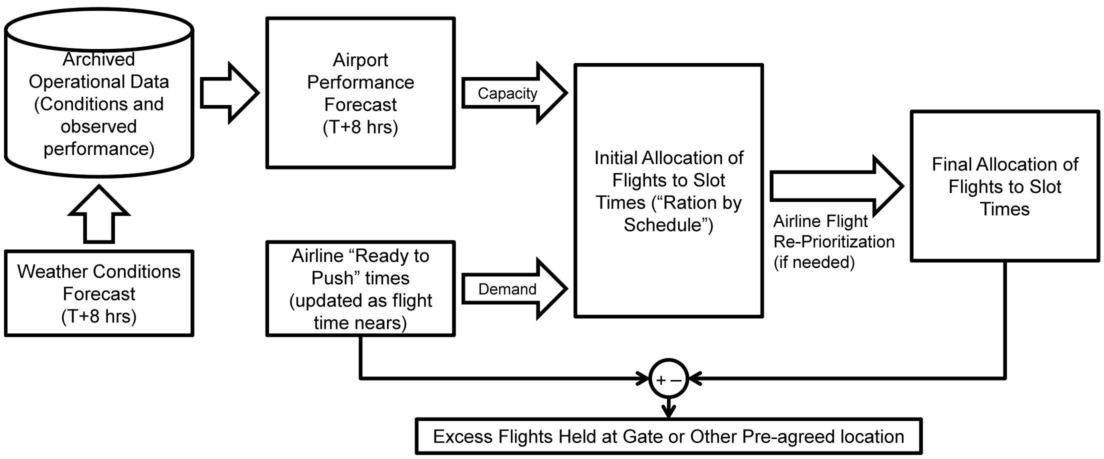
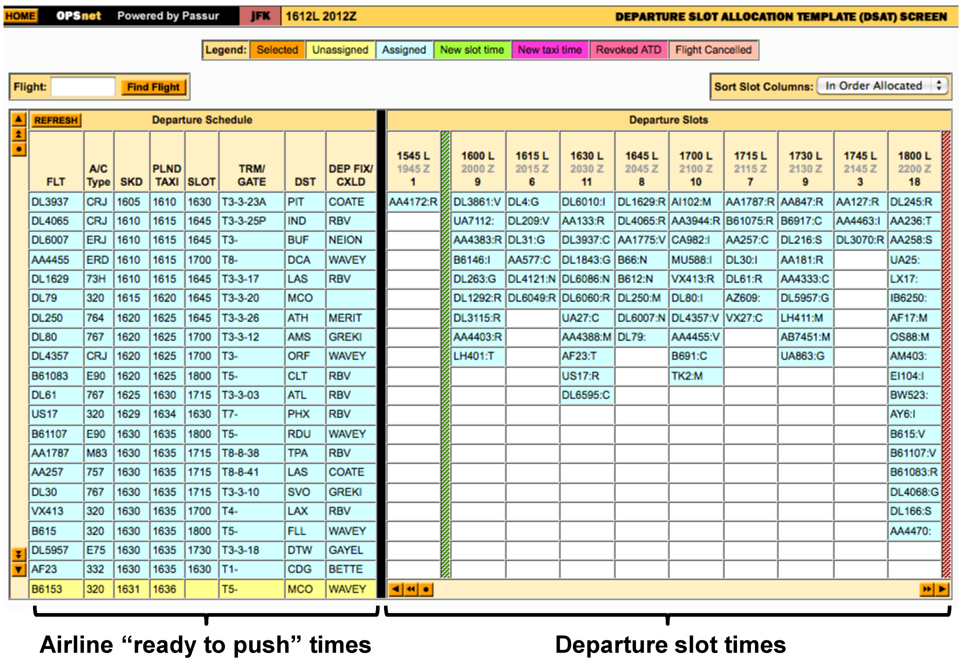



**Figure 13: JFK Surface Congestion Management (SCM) Approach**

Figure 14 shows an example "Departure Slot Allocation" screen from the system employed at JFK. The left side illustrates airline-sourced "ready to push" times by flight, while the right side shows how these flights were allocated to departure slots in 15 minute time bins. The green vertical bar delineates the current time bin. Differences between the "ready to push" and departure slot times represent the gate hold time to manage surface congestion more efficiently. For example, DL1629 had a desired push time of 16:15 but a slot time of 16:45 so received a 30 minute gate hold time and shows as demand in the 16:45 departure slot window.




| FLT | A/C Type | SKD | PLND TAXI | SLOT | TRM/ GATE | DST | DEP FIX/ CXLD |
|---|---|---|---|---|---|---|---|
| DL3937 | CRJ | 1605 | 1610 | 1630 | T3-3-23A | PIT | COATE |
| DL4065 | CRJ | 1610 | 1615 | 1645 | T3-3-25P | IND | RBV |
| DL6007 | ERJ | 1610 | 1615 | 1645 | T3- | BUF | NEION |
| AA4455 | ERD | 1610 | 1615 | 1700 | T8- | DCA | WAVEY |
| DL1629 | 73H | 1610 | 1615 | 1645 | T3-3-17 | LAS | RBV |
| DL79 | 320 | 1615 | 1620 | 1645 | T3-3-20 | MCO | |
| DL250 | 764 | 1620 | 1625 | 1645 | T3-3-26 | ATH | MERIT |
| DL80 | 767 | 1620 | 1625 | 1700 | T3-3-12 | AMS | GREKI |
| DL4357 | CRJ | 1620 | 1625 | 1700 | T3- | ORF | WAVEY |
| B61083 | E90 | 1620 | 1625 | 1800 | T5- | CLT | RBV |
| DL61 | 767 | 1625 | 1630 | 1715 | T3-3-03 | ATL | RBV |
| US17 | 320 | 1629 | 1634 | 1630 | T7- | PHX | RBV |
| B61107 | E90 | 1630 | 1635 | 1800 | T5- | RDU | WAVEY |
| AA1787 | M83 | 1630 | 1635 | 1715 | T8-8-38 | TPA | RBV |
| AA257 | 757 | 1630 | 1635 | 1715 | T8-8-41 | LAS | COATE |
| DL30 | 767 | 1630 | 1635 | 1715 | T3-3-10 | SVO | GREKI |
| VX413 | 320 | 1630 | 1635 | 1700 | T4- | LAX | RBV |
| B615 | 320 | 1630 | 1635 | 1800 | T5- | FLL | WAVEY |
| DL5957 | E75 | 1630 | 1635 | 1730 | T3-3-18 | DTW | GAYEL |
| AF23 | 332 | 1630 | 1635 | 1630 | T1- | CDG | BETTE |
| B6153 | 320 | 1631 | 1636 | | T5- | MCO | WAVEY |

| 1545 L 1945 Z 1 | 1600 L 2000 Z 9 | 1615 L 2015 Z 6 | 1630 L 2030 Z 11 | 1645 L 2045 Z 8 | 1700 L 2100 Z 10 | 1715 L 2115 Z 7 | 1730 L 2130 Z 9 | 1745 L 2145 Z 3 | 1800 L 2200 Z 18 |
|---|---|---|---|---|---|---|---|---|---|
| AA4172:R | DL3861:V | DL4:G | DL6010:I | DL1629:R | AI102:M | AA1787:R | AA847:R | AA127:R | DL245:R |
| | UA7112: | DL209:V | AA133:R | DL4065:R | AA3944:R | B61075:R | B6917:C | AA4463:I | AA236:T |
| | AA4383:R | DL31:G | DL3937:C | AA1775:V | CA982:I | AA257:C | DL216:S | DL3070:R | AA258:S |
| | B6146:I | AA577:C | DL1843:G | B66:N | MU588:I | DL30:I | AA181:R | | UA25: |
| | DL263:G | DL4121:N | DL6086:N | B612:N | VX413:R | DL61:R | AA4333:C | | LX17: |
| | DL1292:R | DL6049:R | DL6060:R | DL250:M | DL80:I | AZ609: | DL5957:G | | IB6250: |
| | DL3115:R | | UA27:C | DL6007:N | DL4357:V | VX27:C | LH411:M | | AF17:M |
| | AA4403:R | | AA4388:M | DL79: | AA4455:V | | AB7451:M | | OS88:M |
| | LH401:T | | AF23:T | | B691:C | | UA863:G | | AM403: |
| | | | US17:R | | TK2:M | | | | EI104:I |
| | | | DL6595:C | | | | | | BW523: |
| | | | | | | | | | AY6:I |
| | | | | | | | | | B615:V |
| | | | | | | | | | B61107:V |
| | | | | | | | | | B61083:R |
| | | | | | | | | | DL4068:G |
| | | | | | | | | | DL166:S |
| | | | | | | | | | AA4470: |

**Figure 14: Airline Ready to Push and Departure Slot Allocation Example [6]**

## 3.2 METHODOLOGY

There are many potential impacts of surface congestion management, for example in terms of taxi-out time, fuel burn, emissions, throughput, gate usage, holding area usage, ground crew operations, passenger connectivity, bag connectivity, airport terminal occupancy, airport terminal revenues, etc. *The focus of the analysis reported here is a first order assessment of annualized impacts of the 2010 surface congestion management approach on taxi-out time, fuel burn and* $CO_2$ *emissions at JFK.* The general approach to achieve this was to compare taxi times, fuel burn and emissions pre/post surface congestion management implementation, with all other relevant operational factors being as equal as possible. It was possible to find a few days where the airport was operating under very similar conditions pre/post surface congestion management implementation, allowing the general impacts of the technique to be observed. For example, Figure 15 shows that, on these sample days, surface congestion management reduced the number of aircraft on the airport surface between 17:00 and 21:00 (corresponding to the evening departure push at JFK) from a peak of 40 on the sample day in the period before the technique was implemented to about 25 after it was implemented, resulting in active taxi-out time savings of over 20 minutes for the average flight departing at 20:00. The surface traffic snapshot shown in Figure 16 reinforces the effect in terms of the reduced departure queue size and resultant reduced taxi-out times, with the "excess" aircraft being held off the active movement area.

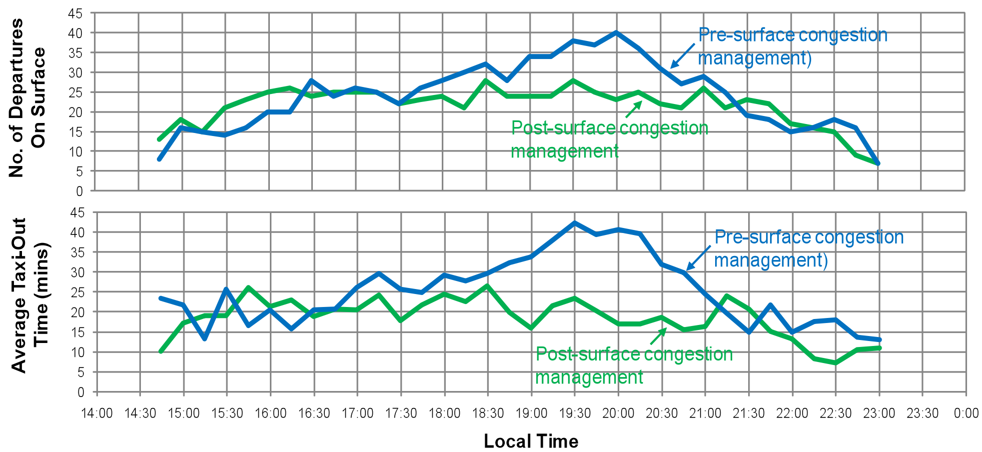



**Figure 15: Comparison of Taxi-out Times Pre/Post Surface Congestion Management Implementation for Sample Days**

Although these observations provide insights into the effect of surface congestion management, data across numerous days is required to estimate annualized impacts. However, the large number of factors that influence airport operations (e.g., demand, capacity, airport configuration, weather/ATC constraints, equipment status, etc.) and the complexity of operations specifically at JFK made finding a large enough sample of comparable days pre/post-implementation very difficult. Therefore, an analysis approach was developed which found relationships between surface congestion management and taxi time impacts in each major airport configuration and then applied the identified relationships to the full set of data to determine the annualized impacts of the congestion management technique, as described in the next section.

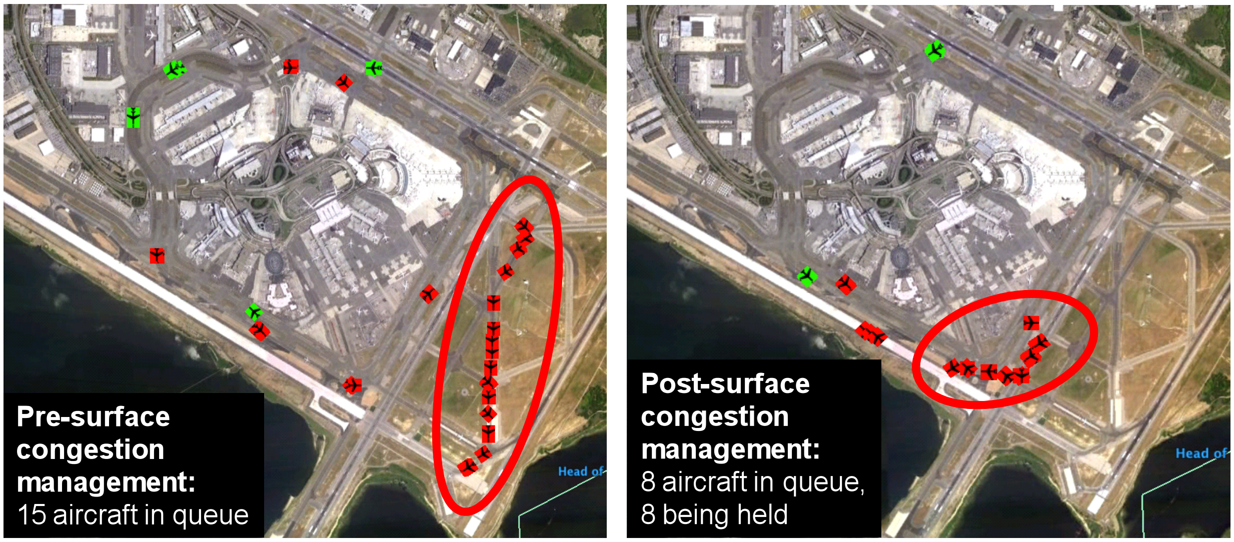



**Figure 16: Comparison of Airport Traffic and Departure Queues Pre/Post Surface Congestion Management**

The analysis methodology is presented in Figure 17 with the general sequence of steps presented along the top and more detail on how the steps were executed below. Each of the steps is discussed in detail in this section.

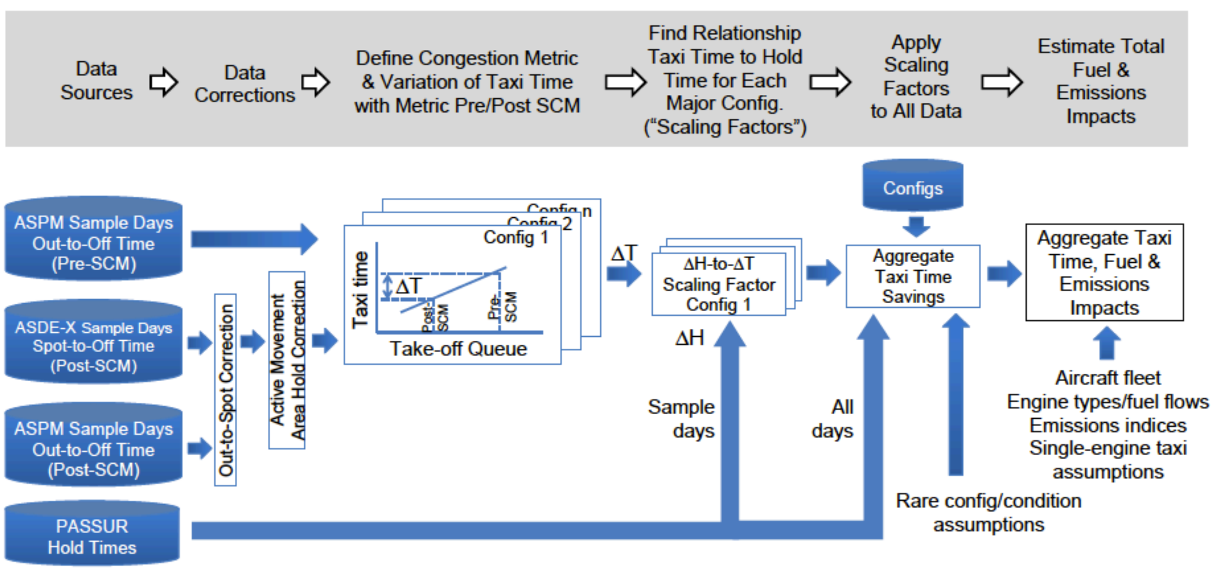



**Figure 17: Analysis Methodology**

### 3.2.1 Data Sources

This analysis used the ASPM database [1] which provides flight-specific OOOI times and airport throughput in 15 minute intervals; ASDE-X data which provides position in the active movement area (not ramp) at 1 second updates; and the PASSUR program data which provides flight-specific desired and slot times.

The pre-implementation analysis period was selected to be January 1, 2009 - December 31, 2009. The initiation of the surface congestion management process coincided with the closure of runway 13R/31L, but the impacts during the runway closure were not analyzed because the airport was not in its normal state (i.e., there was no pre-implementation data corresponding to JFK without runway 13R/31L). Therefore, the post-implementation analysis period was selected to be July 1, 2010 - December 31, 2010 corresponding to the day runway 13R/31L re-opened through the last day for which all of the data sources discussed above were available for this analysis.

### 3.2.2 Data Corrections

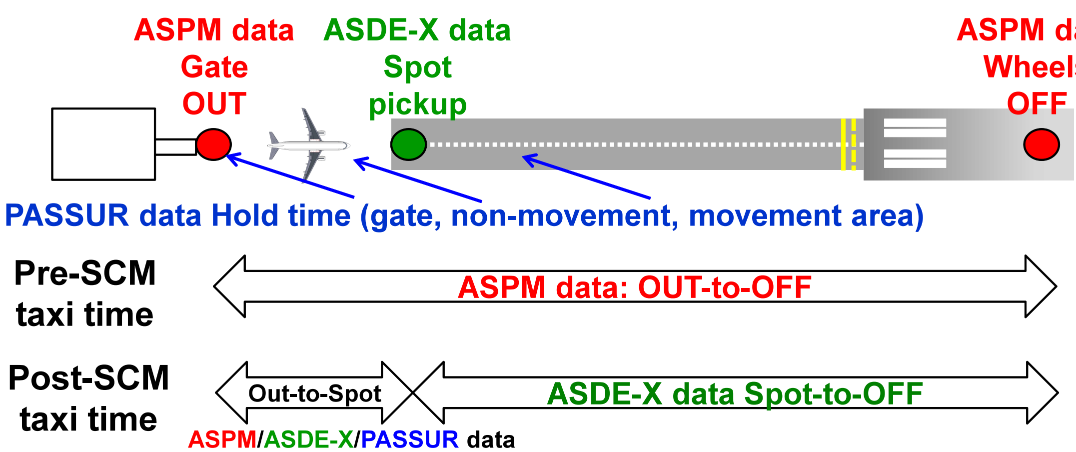



**Figure 18: Key Analysis Events**

The data sources identified above provided the key analysis events illustrated in Figure 18. The difference between the OOOI OUT and OFF times provided a good measure of the taxi-out time in the pre-surface congestion management environment. However, it was not suitable in the post-surface congestion management environment due to the fact that a large number of the flights which were given slot times after their desired push times were held "off-gate". In those cases, the ASPM OUT time was not an accurate reflection of when the aircraft actually started taxiing to its departure runway, but rather when it left the gate to be held elsewhere (as would happen if the gate was needed for an inbound arrival). Therefore, the post-surface congestion management taxi-out times were determined from the ASDE-X data. Given that the ASDE-X tracks were generally picked up at the spots (the interface between the ramp and active movement areas) the tracks needed to be corrected back to an equivalent OUT time so they could be directly compared to the pre-implementation taxi-out times based on the OUT-to-OFF events.

To determine the appropriate OUT-to-spot correction factor, distributions of the differences between ASPM OUT times and ASDE-X pickup times were calculated for pre- and post-surface congestion management days. For the pre-implementation case, only 9 weeks of ASDE-X data were available, whereas 6 months of data were available for the post-implementation period. The pre- and post-implementation distributions were subtracted from each other resulting in the left side of Figure 19, which shows a spike above the horizontal axis and a trailing tail below it. The positive spike represents additional flights pre-congestion management implementation with small differences between their out and pickup times, while the trailing tail represents additional flights post-surface congestion management implementation with large differences. Because the number of flights in the negative tail and positive spike is approximately equal, it was hypothesized that the trailing tail represent flights that, pre-implementation, pushed back normally but post-implementation were held off-gate (resulting in a long period of time between their OUT and spot times). The positive spike therefore represents a distribution of typical OUT-to-spot times. This subsequently had a normal distribution fitted to it as shown on the right side of Figure 9, with a resulting mean of 7 minutes and a standard deviation of 2 minutes. This can be interpreted as the distribution of times it takes a typical flight at JFK to reach the spot once the parking brake has been released, accounting for tug push-back, engine start and checklist completion times.

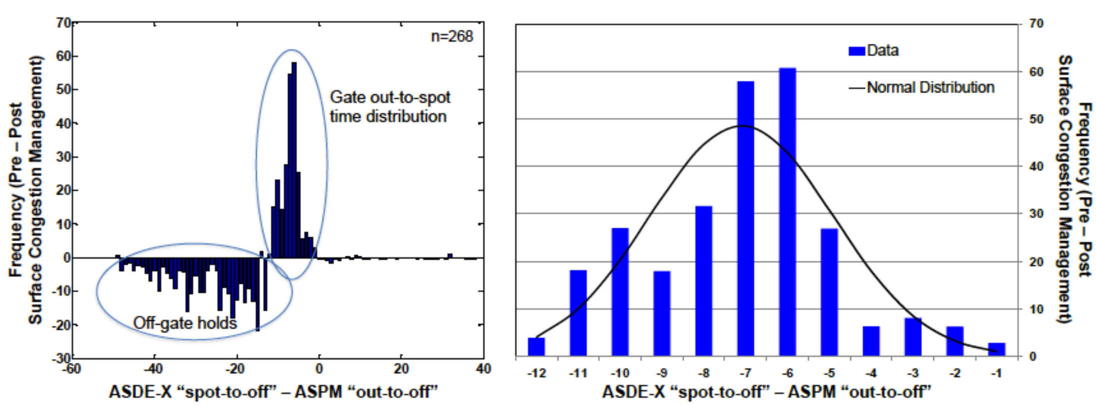



**Figure 19: "OUT-to-spot" Correction Factor Data**

Another correction factor was required to account for those flights that held in the active movement area at pre-designated hold locations until their slot time; i.e., their ASDE-X pick-up time was not a true reflection of their start of taxi time (similar to the reason why ASPM data was not appropriate for any flight with an off-gate hold). To correct for the fact that these flights were in fact holding in the active movement area (most likely with engines off), all flights which appeared in ASDE-X data 7 minutes or more before their scheduled slot time (5 minute PASSUR allowance + 2 minute grace period) had their spot times moved forward to their scheduled slot time. This approach was validated by examining ASDE-X tracks for individual flights that fit the criteria and verifying that those flights stayed in their assigned hold area until their slot time, and then began taxiing to their departure runway.

### 3.2.3 Define Congestion Metric & Variation of Taxi Time with Metric pre/post SCM

The key congestion metric used in this analysis was the "take-off queue", introduced by Idris et al. [16], which for a given flight $i$ is defined as the number of other take-offs which occur between the pushback and take-off time of aircraft $i$. Other metrics were also tested, including number of departing aircraft on the airport surface and the number of aircraft in physical departure queues at the runways, but they were found to be less suitable for JFK analysis. The main advantage of the take-off queue versus the number on surface is that it takes overtakes into account, when an aircraft takes off before another aircraft that left its gate before the first aircraft did. The geometry of JFK results in many such flights because some gates are much closer to the departure runway than others. As a result, the take-off queue provides a much better estimate of the time that a flight will spend taxiing. The downside of this metric is that it is flight-specific and more difficult to calculate. Figure 20 shows that the take-off queue is a better predictor of taxi time at JFK as measured by $R^2$.

To convert the change in take-off queues into a change in taxi time at JFK, a regression was calculated using taxi time versus take-off queue data as shown in Figure 21. The slope of the regression can be interpreted as the incremental taxi time for every additional aircraft in the take-off queue. The slopes of the regression pre- and post-surface congestion management are very similar, indicating the dynamics of the airport are unaffected by the procedure, but the airport is operating at much lower average take-off queue counts when surface congestion management is in operation. Regressions like this were calculated for the top six most common configurations that experienced holds at JFK and the regression line slopes of all but one of the configurations statistically equal pre- and post-implementation, but did vary between configurations as expected given their different capacities.

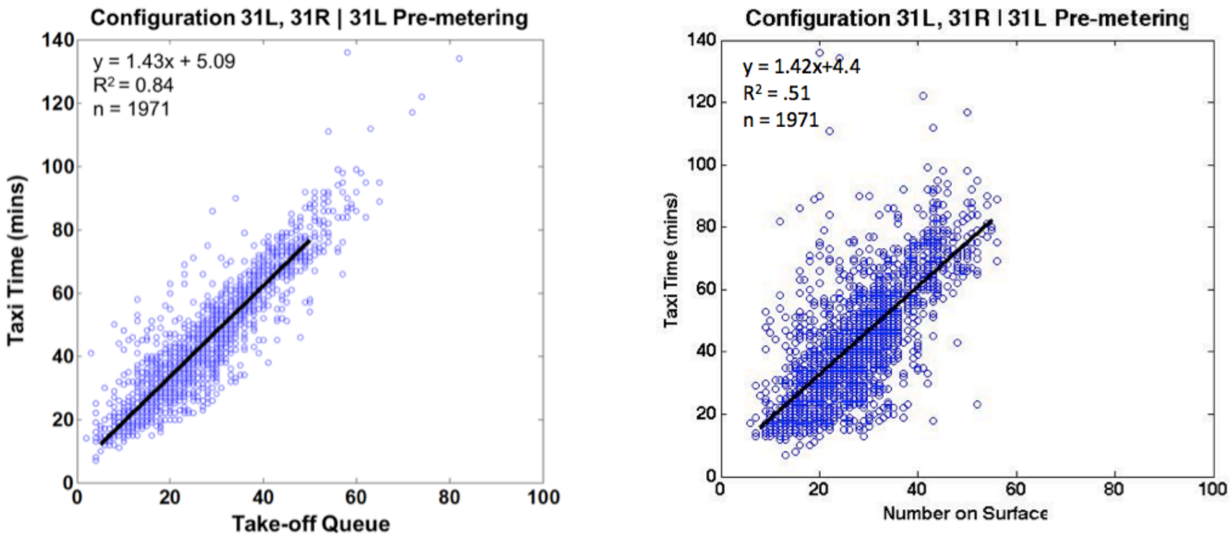



**Figure 20: Comparison of Traffic Metrics**

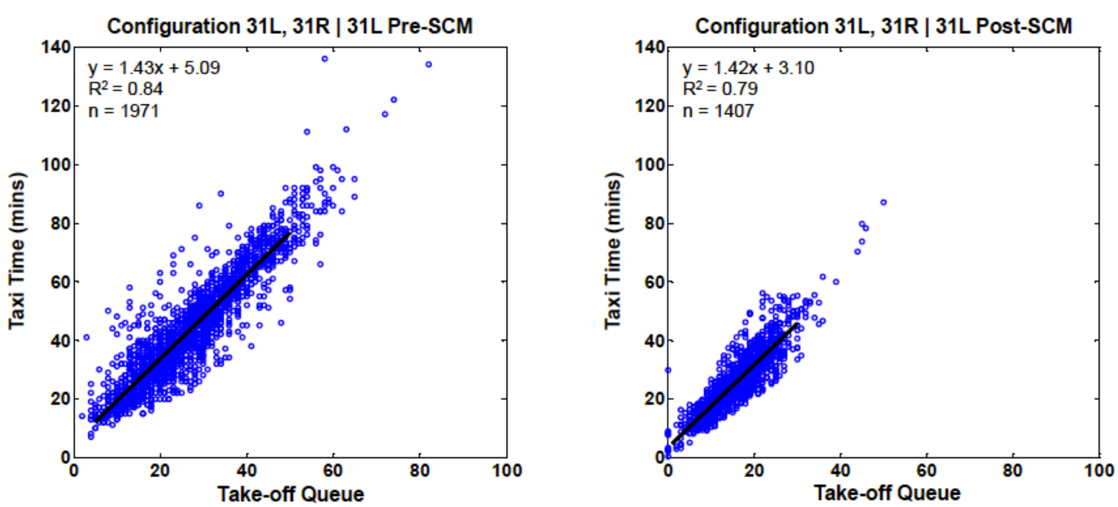



**Figure 21: Relationship of Taxi Time to Take-off Queue**

To alleviate the problem highlighted earlier with identifying similar days pre- and post-surface congestion management implementation, multiple "sample days" were found for each of the top 6 configurations. These days were chosen by looking at the peak departure period (5-9 PM in most cases) and finding days when the airport stayed in the same configuration for the duration of the period. This eliminated instances where the configuration was changed midway through the period, which could affect the results. By looking at a group of days and averaging the traffic over them, the variations in operation from day to day are accounted for to first order. The average takeoff queue across the group of sample days was calculated in 15 minute bins (e.g., 17:00-17:15) pre- and post-surface congestion management implementation, and using the regression lines for each configuration, the taxi time impact of the technique was determined in those 15 minute time bins. This was then summed over all time periods in the sample days to determine a total amount of taxi time saved.

### 3.2.4 Find Relationship of Taxi Time to Hold Time for Each Major Configuration ("Scaling Factors")

The difference in taxi time observed from the previous step can be compared to the hold time (defined as the difference between the desired push time and the slot time) due to the surface congestion management technique to determine configuration-specific "Scaling Factors": see Figure 22.

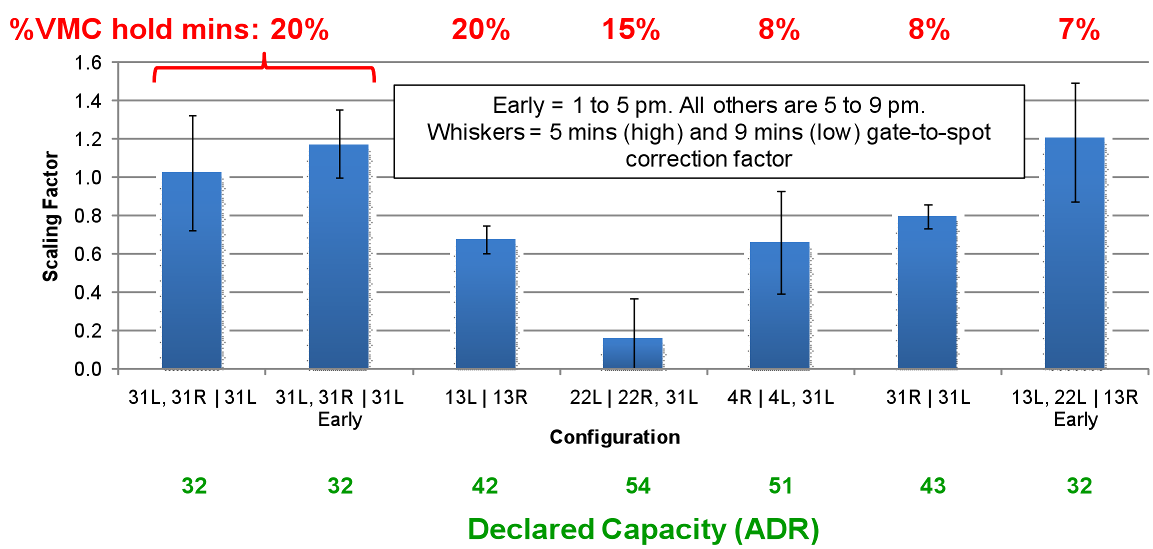



**Figure 22: Scaling Factors Relating Taxi Time Reduction to Hold Time By Configuration**

They can be considered as representing the observed taxi time reduction of each minute of hold time. Notice the sensitivity of the scaling factors (represented by the whiskers) to the gate-to-spot correction factor described earlier. The reasons for the differences between configurations are complex and are a subject of on-going investigation, but there are several hypotheses. The scaling factors appear highest for the configurations with the lowest declared capacity. Because these configurations can accommodate fewer aircraft, it makes sense that for a given demand (and hold) level, there would be more congestion and therefore more benefit from SCM when compared to a higher-capacity configuration. There could also be configuration-specific operational restrictions and policies that could lead to more or less congestion, such as the number of departure fixes available or use of certain configurations only at certain times or weather conditions.

### 3.2.5 Apply Scaling Factors to All Data

Once scaling factors for the main configurations were calculated, they were generalized to the other configurations in use at JFK by comparing the number of runways in use as well as the specific runways (resultant average 0.79). Ideally, a separate analysis would be conducted for IMC conditions. However, because IMC conditions occur infrequently (< 10%) at JFK, there was not enough data to perform a valid analysis. Therefore, the conservative assumption was made that the scaling factors were the same for VMC as IMC for a given configuration. This is a conservative because capacities are generally lower in IMC and hence the benefits of surface congestion management would be larger. This full list of scaling factors was then applied to ALL the gate holds in the six month analysis period to estimate the aggregate taxi time impacts of surface congestion management. This number was doubled to estimate the annualized impacts.

### 3.2.6 Estimate Total Fuel & Emissions Impacts

To convert from taxi time savings into fuel and emissions savings, an average fuel burn index was calculated for each month of the study period to account for changes in fleet mix. The PASSUR data included the tail number of all aircraft. A fleet database was used to match tail numbers to engine types, and ICAO ground idle fuel flow certification data [17] was used to estimate the taxi fuel flow rate for each aircraft accounting for the number of engines of each type it possessed and APU/single-engine taxi assumptions. Fuel burn savings from surface congestion management were determined by multiplying this fuel flow rate by the taxi time savings determined from the previous steps and summing over all flights. Fuel burn savings were converted to carbon dioxide emissions savings by using the standard $CO_2$ emissions index of 3.16 kg $CO_2$/kg fuel burnt.

## 3.3 RESULTS

### 3.3.1 Key Results

Table 1 presents the calculated impacts of surface congestion management at JFK once the methodology discussed above has been applied. Total annualized taxi time reductions of 14,800 hours translate into annual savings of 5.0 million gallons of fuel and 48,000 metric tons of carbon dioxide from surface congestion management at JFK. The total taxi time reduction results are within the range of estimates from *simulation* studies in the open literature [13,14,15], but the results shown in Table 1 are based on the actual *operational* data. At JFK in 2009 there were 104,000 total hours spent taxiing out [1], which corresponds to total fuel burn of 35.2 million gallons of fuel using the methodology from Section 3.2.6. Therefore, the SCM program had savings of 14% over the current surface operations in terms of fuel and time saved. Taking the BTS estimate of 16.2 billion gallons of fuel consumed by certificated carriers in the US in 2009 [18] and scaling by the number of flights at JFK relative to the NAS, we can make a rough estimate of total fuel burn (all phases of flight) for departures at JFK to be 369 million gallons, making the savings from SCM 1.3% of the total. The estimate of 1.3% is probably an overestimate because the BTS estimate only includes US carriers. Including international carriers could drop the the estimate to 1% or lower. The ASPM estimate has no such caveat.

**Table 1: Calculated Benefits**

| Configuration | Proportion of Hold Mins | Hold Time ($10^4$ mins) | Scaling Factor | Taxi-out Time Reduction ($10^4$ mins) | Fuel Reduction (US gallons) | Carbon Dioxide Reduction (metric tons) |
|---|---|---|---|---|---|---|
| 31L,31R \| 31L | 20% | 11.8 | 1.17 & 1.02 | 13 | 730,000 | 6,990 |
| 13L \| 13R | 18% | 10.4 | 0.67 | 7 | 391,000 | 3,750 |
| 22L \| 22R, 31L | 13% | 7.5 | 0.16 | 1.2 | 66,200 | 630 |
| 4R \| 4L, 31L | 9% | 5.2 | 0.66 | 3.4 | 191,000 | 1,830 |
| 31R \| 31L | 7% | 4.3 | 0.79 | 3.4 | 187,000 | 1,790 |
| 13L, 22L \| 13R | 6% | 3.5 | 1.2 | 4.2 | 239,000 | 2,290 |
| Others | 27% | 15.5 | 0.79 | 12.2 | 690,000 | 6,600 |
| **Totals (6 months)** | | **58.1** | | **44.4** | **2,490,000** | **23,900** |
| **Totals (annual)** | | | | **88.8 (14,800 hrs)** | **4,980,000** | **47,800** |

Figure 12 shows how the resulting fuel cost savings of surface congestion management at JFK vary as a function of assumed fuel price and percent use of single engine taxi (a taxi procedure where only one engine is turned on: the fuel burn of a single-engine taxi was estimated to be 60% of the equivalent "all engine" taxi). At the typical 2010 fuel price range of $2-3/gallon [18], fuel costs savings through surface congestion management are estimated to be $10-15 million per year at JFK if it is assumed no flights are performing single-engine taxi, and $7.5-12.5 million if half of the flights are assumed to be performing single-engine taxi.

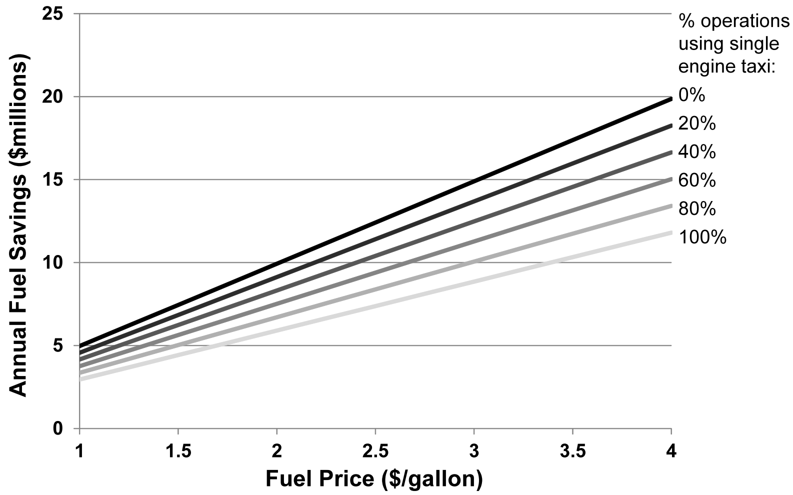



**Figure 23: Annual Monetized Benefits of JFK Metering**

### 3.3.2 Taxi-in Times

One possible side effect of surface congestion management can be an increase in taxi-in times for arriving aircraft if the procedure for holding departure aircraft is not sufficiently well planned. For example, if there are multiple aircraft being held at their gates past their desired departure times, there might not be enough gates available for arriving aircraft, resulting in the arriving aircraft having to wait on the surface and delaying their IN times. The average taxi-in times by hour from 17:00 to 21:00 local time were taken from ASPM for the years 2007 to 2010. Figure 13 shows the average over the entire period for each year. If we propose two hypotheses $H_0$ and $H_a$:

$H_0$: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_a$: not all $\mu_i$ are equal

where $\mu_i$ is the average taxi time in a given year i (i = 1:4 for 2007:2010), then we can use a one-way ANOVA test to test which hypothesis should be chosen. To conclude $H_a$ with an alpha of 0.1, F* must be greater than F(.9, 3,16) = 2.46. We calculate $F^* = \frac{MSTR}{MSE} = 0.117$ where MSTR is the treatment mean square and MSE is the error mean square. We can see that F* is much less than F and that $H_0$ holds, meaning that there is no significant change in the mean taxi-in time. As a result we can conclude that to the first order departure metering has no impact on taxi-in times at JFK.

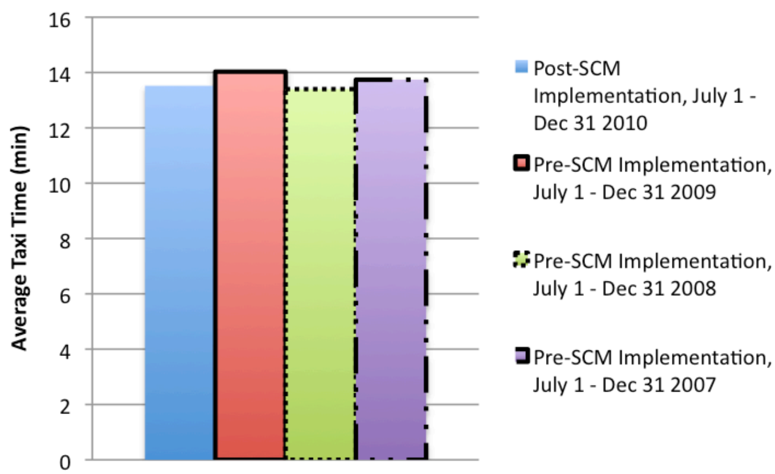



**Figure 24: JFK Taxi-in Time Analysis**

### 3.3.3 Throughput

Another possible side effect of surface congestion management can be reduced throughput if too many aircraft are held back for too long. Figure 14 shows a comparison of airport throughput before and after surface congestion management by configuration and airport-wide. The throughput here is measured by the number of wheels-off times in a given hour. The time period studied is the 5-9 PM period (except for 13L, 22L | 13R, which is 1-5 PM as in the study). The only configuration with a statistically significant (90 % confidence) change was 13L | 13R, which increased after metering was implemented. Overall, there was no change in the average throughput. As a result, we can conclude that to the first order, departure metering has no impact on throughput at JFK. While the increase in throughput in 13L | 13R is significant, it is small and could be due to other factors, such as the improvements on 13R in the first half of 2010.

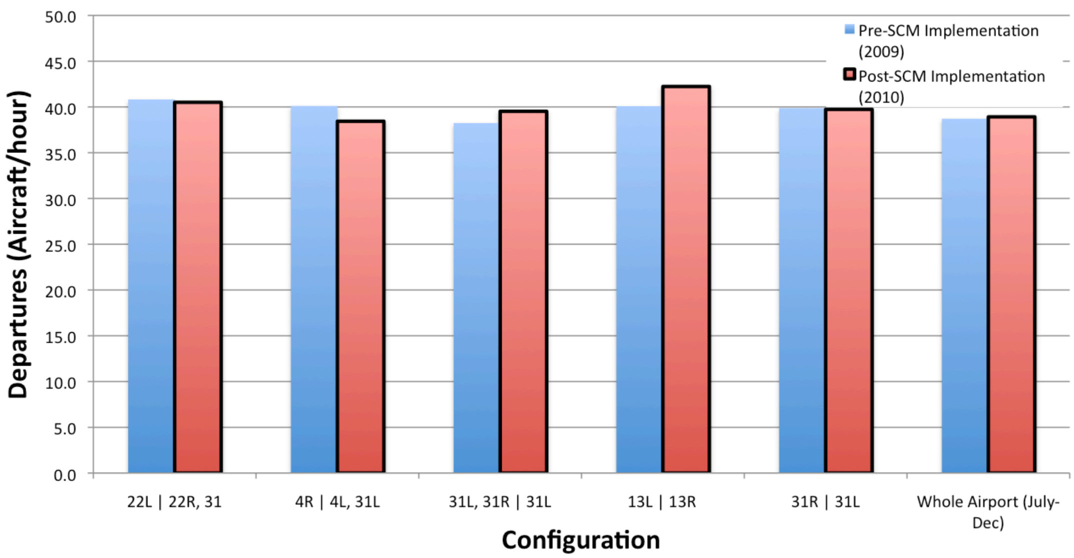



**Figure 25: JFK Throughput Analysis**

## 3.4 COMPARISON OF HIGH FIDELITY ANALYSIS TO SATURATION CURVE METHOD

Benefits from metering at JFK can also be calculated using the saturation curve method introduced in Chapter 1. Figure 26 shows the saturation curve for JFK airport across all configurations and weather conditions in 2009, the last full year before metering was implemented. The saturation point $N^*$ occurs around $N(t) = 25$ and the corresponding saturation throughput is just above 10 departures / 15 minutes. Although N-control was not used to perform metering, this thesis argues that it is a valid method for estimating the possible benefits at all airports of implementing metering using any method. This is because using a saturation curve defines the available pool of benefits by quantifying the inefficiencies that can be addressed by metering. In addition, each different implementation of metering essentially limits the number of aircraft on the surface by controlling the length of different queues.

To compare the estimated benefits of the PASSUR trial with the saturation curve method, an equivalent $N_{Ctl}$ value must be identified. This value represents the aggressiveness of the metering approach. While a value of $N_{Ctl} = N^*$ will be used in the medium fidelity methodology, the value can vary. Examining the ASDE-X data (corrected to include the ramp area and exclude off-gate holds) for the sample days during the post-metering period, the average *N(t)* across all configurations was 18 during the peak traffic hours. This was taken as the equivalent $N_{Ctl}$. While this seems low, especially given N* = 23 for the previous year, it can be viewed as an upper bound to the benefits. In addition, because N-control is not being used there is no clear cap on N, necessitating the use of the average that results in a more aggressive value. Finally, there is anecdotal evidence that the metering program at JFK was aggressive in reducing congestion, which would correspond to a lower $N_{Ctl}$.

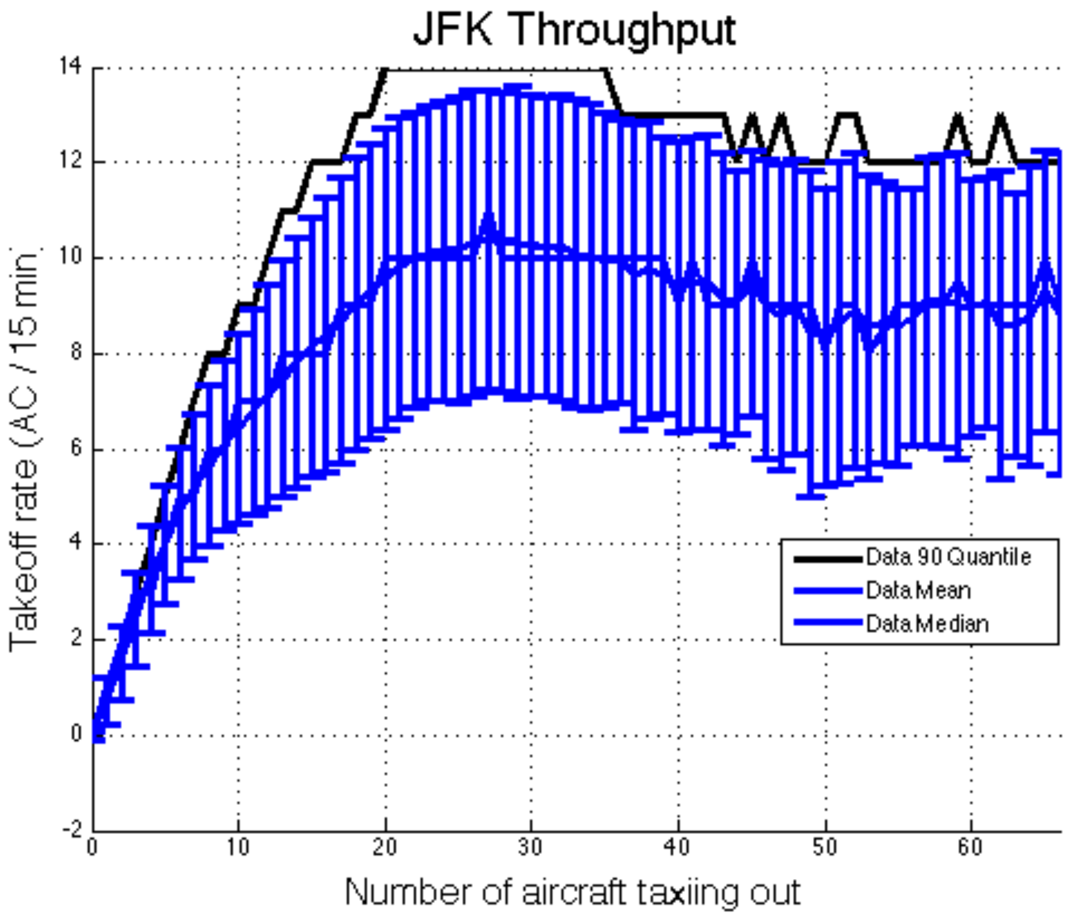



**Figure 26: JFK Saturation Curve - 2009**

Returning to the pre-metering ASPM data, the total number of flights operating when *N(t)* > 18 was calculated to be 73,166, with an average taxi time of 46.0 minutes. The average taxi time of flights with *N(t)* = 18 was 31.2 minutes. Therefore, the benefits from moving all flights in congestion to the control point would be

$$\text{Benefits} = \text{Flights}_{\text{congested}} * (\text{AvgTaxiTime}_{\text{congested}} - \text{AvgTaxiTime}_{\text{saturation}})$$

$$= 73{,}166*(46-31.2) = 108{,}000 \text{ minutes} = 18{,}000 \text{ hours}$$

The benefits for metering to N = 23 are 12,260 hours, resulting in 2 values that bracket the observed value of 14,800 hours. The first estimate is higher than the observed benefits, but is close to other studies' estimates of the benefits of metering [11,12]. In addition, it represents the theoretical benefits. As we have shown, there are a number of real-world constraints such as off-gate holds, the use of 15 minute bins, and user compliance that could reduce the achieved benefits.

While N(t) = 18 is below N*, which would indicate that the full throughput of the airport is not being achieved, it is important to remember that this was not the method used at JFK and is instead a way of further verifying the results shown in this section. The throughput was shown in section 3.3.3 to not have significantly changed from 2009 to 2010.

## 3.5 CONCLUSIONS

The analysis of the field trial at JFK provides useful insights for the extension of the benefits assessment into the future. The differences between benefits in specific configurations is not immediately intuitive but does make sense given that taxi distance, taxi time and airport throughput can all vary significantly based on the configuration. As will be explained in the next chapter, performance and congestion in the future are calculated by configuration (although the issue of changing airport behaviour and configuration choice is not examined). Our assertion that saturation curves can describe all types of metering was further supported by the comparison between the results from the high fidelity analysis and the saturation curve method. Finally, the overall efficiency of metering at JFK in terms of reduction in taxi time to hold time is relatively large (overall scaling factor of 0.79) despite a substantial amount of off-gate holds and large differences in performance by configuration. In addition, possible secondary effects of metering such as reduced throughput and increased taxi-in times were shown to be negligible. Given these findings, secondary effects will be neglected in the medium and low fidelity analyses, as will the efficiency of metering (scaling factors). However, landside constraints such as gate availability (the cause of off-gate holds) will be considered.

# 4. MEDIUM FIDELITY METHOD FOR ASSESSING SCM BENEFITS AT MULTIPLE AIRPORTS

As was discussed in Chapter 1, there is a need to perform benefits assessments on potential air traffic management techniques including SCM. Current benefits estimates are based on operational data that are not available for future years. In addition, traffic levels are estimated to rise to levels never before seen at many airports, meaning that the behavior of these airports could be different from what has previously been experienced.

A simple formulation of the problem is as follows: the benefits from metering at a generic airport in the future must be calculated. The inputs available are demand estimates in the form of estimated pushback schedules in 5 year increments and estimated changes in capacity due to construction or ATC technology improvements in the same increments, and operational data from the previous 10 years is available from ASPM.

This chapter presents a medium fidelity methodology to calculate future benefits. It is informed by the high fidelity analysis of JFK airport in the previous chapter but has a wider applicability and uses more generic techniques. While it does not account for most of the secondary effects studied in the JFK model, it still examines the chosen airports in detail by simulating taxi times and calculating saturation curves specific to an airport and a configuration. The results from this method will be compared to both historical data and the field trial results from JFK and BOS. In addition, they will be used to construct NAS-wide estimates in the low fidelity model.

## 4.1 BACKGROUND

Of the methodologies described in Chapter 2, most are not appropriate for the task described above due to the nature of the future data available. The N-control approach of using saturation curves to quantify congestion was chosen for its generalizability and adaptability. If take-off times can be calculated from the future pushback schedules, saturation curves can be derived for the future that allow calculation of the congestion at a given airport. By adapting the taxi time simulation developed by Simaiakis [19], characteristics of specific airports such as configuration and layout are replaced by general values such as average unimpeded time and maximum throughput.

The work performed by Sensis [13,14,15] relied heavily on ASDE-X data and simulations. This level of detail is not available for the future, and while simulations could be developed from present-day data there are several formidable issues. Each simulation would have to be tailored to a specific airport because it calculates the exact path each flight takes, requiring a great deal of time and familiarity. In addition, ASDE-X data is not readily available for many of the desired study airports. Finally, new construction at an airport would require the simulation to be changed, but the way the new infrastructure will be used is highly uncertain.

The CDQM approach of setting a target delay has already been discussed and shown to be similar to the saturation curve method in Chapter 2.

The remainder of this chapter will describe both the techniques needed to develop the data to construct and analyze future saturation curves and the results from applying those techniques at 8 airports. Section 4.2 will set out the complete methodology used to calculate future taxi times and construct saturation curves. In addition, it will describe how physical constraints were taken into account. Section 4.3 introduces the airports studied, and Section 4.4 presents the results from applying the methodology as well as comparisons to the high-fidelity model from JFK, the field trial at BOS, and historical benefits generated from ASPM-based saturation curves.

This methodology is necessarily at a lower fidelity than the one done for the case study of JFK. Secondary effects such as single-engine taxi and off-gate holds are not examined in relation to benefits, and because the saturation curve method does not explicitly simulate metering (it only identifies the opportunity for savings), the configuration-specific relationship of hold time to saved taxi time is also not examined. In addition to the constraints of the simulation, forecasts are notoriously unreliable. As has been discussed earlier, the benefits of SCM can be highly sensitive to small changes in demand at airports that are near capacity. The secondary effects studied in the high-fidelity methodology are negligible compared to the effects of changing the forecast demand or capacity. This medium fidelity model does still provide a fair amount of customization to individual airports and will be shown to match well with field trial data.

## 4.2 ANALYSIS METHODOLOGY

### 4.2.1 Simulation

The high-level methodology for the benefits assessment of SCM is illustrated in Figure 27.

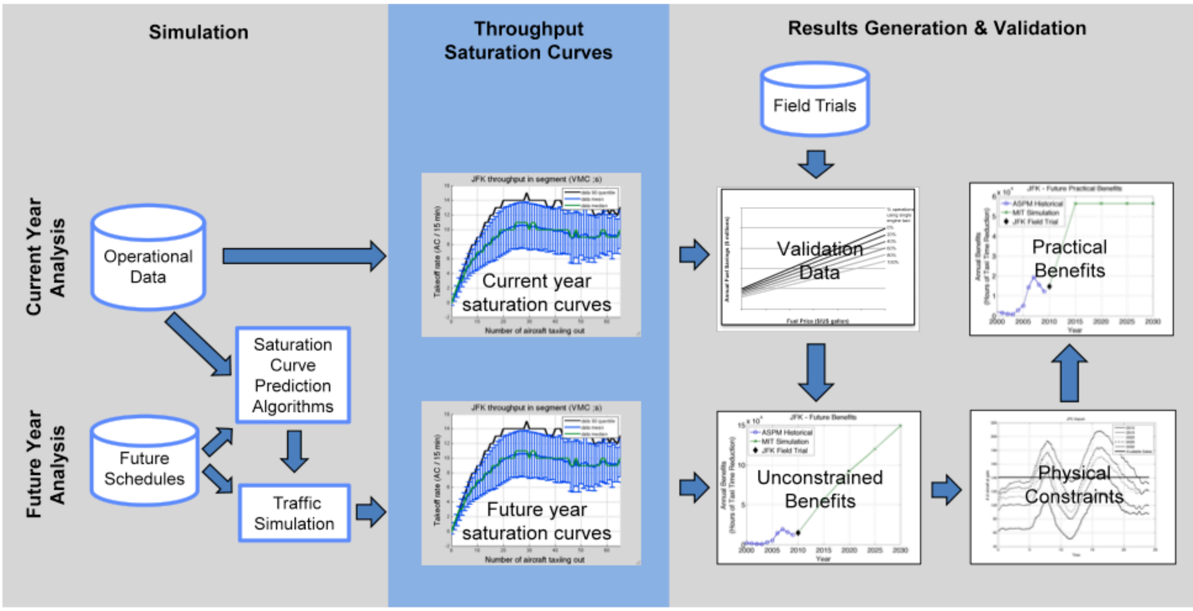



**Figure 27: Departure Metering Analysis Methodology**

The methodology is split into 2 sections that are linked by the concept of saturation curves. The simulation section takes as inputs current operational data to train the prediction algorithms and traffic simulation, and the future schedule data that is used in those simulations once they have been created. The simulation outputs saturation curves for the future year cases (including 2010 as validation).

The congestion and corresponding benefits from SCM are calculated from the saturation curves as well as compared to the field trials discussed in Chapter 2. As will be explained, these results are 'unconstrained' due to the nature of the model behind the future year schedules. The physical constraints on the benefits (gate utilization) are examined and applied to the results to obtain practical benefits levels.

Future year benefits were calculated by simulating throughput saturation curves and congestion at each study airport for the future "out-years" of 2015, 2020, 2025, and 2030 as well as the "current year" 2010. This required the development of a two-stage model that predicted the future saturation curves (Saturation Curve Prediction Algorithm) as well as the future traffic and congestion (Traffic Simulation) to determine where the study airports were operating along these curves in different years.

### 4.2.2 Saturation Curve Prediction

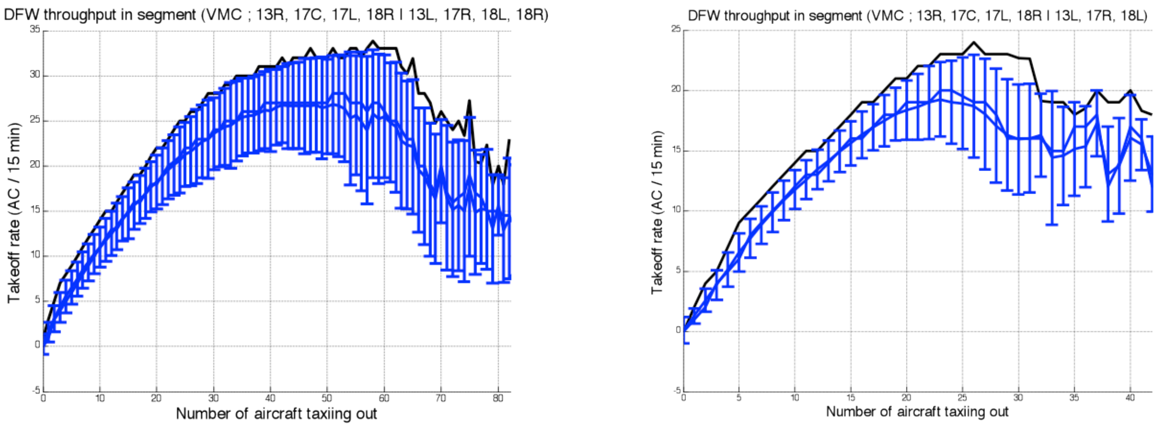



**Figure 28: DFW Saturation Curves in 2000 (Left) and 2010 (Right)**

At first glance, one might assume that a saturation curve is an unchanging characteristic of a given configuration. After all, they are a way of representing throughput, and the maximum achievable throughput should be the declared capacity. This would be a naïve approach, however. In the left side of Figure 28 the average saturation throughput can be seen to be approximately 27 aircraft / 15 minutes, while the average declared capacity for the year was 29 aircraft / 15 minutes. While both these values can vary (and do, as shown by the whiskers) based on factors like the number of arrivals and the downstream weather among others, on average the saturation throughput is below the declared capacity. This is because the saturation throughput better reflects the sustainable capacity. If an airport operates at high demand levels for a period of several hours, the declared capacity cannot be sustained due to uncertainty, delays, varying fleet mix, and a variety of other reasons.

In addition, saturation throughput in the same configuration can change over time as shown by Figure 28 (The omission of 18R in 2010 is irrelevant because simultaneous departures on 18L and 18R are not possible). Our hypothesis is that differing levels of demand are the main driver behind this change. Figure 29 shows how the demand, declared capacity, saturation throughput and taxi times at DFW vary from 2000 to 2010. DFW was at its highest level of demand in 2000 before decreasing constantly (besides a brief spike in 2004). Correspondingly, the saturation throughput was highest between 2000 and 2002 and also decreased thereafter. According to our hypothesis, the decrease in saturation throughput from 2003 onwards happened because the pressures to maximize throughput were removed. The time scale of the effect appears to be 2 or 3 years, as the saturation throughput did not immediately follow the decrease in demand in 2001 and did not rebound in 2004. The declared capacity remained higher for longer (until 2005) perhaps because of institutional memory, before decreasing. Note that even though the saturation capacity decreased, the taxi times (which can be viewed as the amount of pressure on the system) stayed constant or decreased as well. There are other hypotheses on the change in DFW performance examined in Section 4.6.1. Because DFW experiences such a large change in performance a combination of factors is likely responsible, with many of them being specific to the airport.

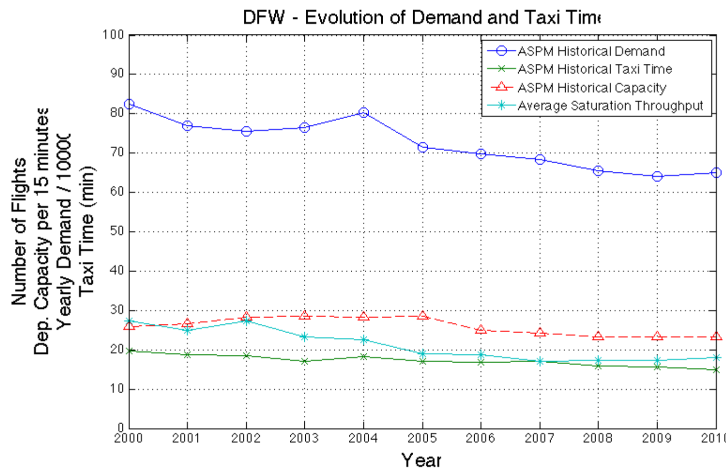



**Figure 29: DFW Demand and Capacity**

Given this complex relationship between saturation capacity and demand and declared capacity, saturation curves needed to be modeled in the future. In addition, it was hypothesized that they could depend on other airport variables such as the number of runways in use or the percentage of capacity used, especially when an airport is expected to construct a new runway (creating a new configuration). Finally, a linear model is not very appropriate because while saturation throughput could significantly decrease as seen at DFW, there is an upper limit to the amount it can increase that is determined by separation requirements between flights, which a linear model would not capture.

Instead, future saturation curves were estimated using the Random Forest (RF) method [20]. The Random Forest was chosen because of the many parameters and conditions that affect an airport's performance, as well as the non-linearity of the performance. The RF method uses groups of decision trees that test the importance of different parameters in order to predict values by calculating the average over all predictions from the individual trees. RF is appropriate for our analysis because it makes no assumptions about the functional relationship between the input/predictor variables and the output, and avoids biases by not assuming a particular function is the correct form to describe airport behavior.

The saturation point and saturation throughput are the target prediction variables that define the airport throughput saturation curves to first order. The input parameters to the random forest model were chosen using engineering judgment as well as the input of subject matter experts and included the mean and peak hourly demand and capacity, the usage of a configuration, the physical size of the airport, and the number of gates. The decision trees were trained, or 'grown', on data from 2000 to 2010, those being the years for which ASPM data exists. Data based on the capacity growth forecasts and future schedules, supplemented with parametric variation of the curves as appropriate for representative days/conditions for the future study years, were input into the model to obtain the future saturation curves.

The saturation point is defined for the purposes of calculation as the first point at which the throughput reaches 95% of its maximum value. To eliminate the high variability due to small sample sizes (outliers at high N values with abnormally high throughput), the 2% of data with the highest N values were removed from the data set for the calculation of N*. The saturation throughput is simply the mean throughput at the saturation point. While the 95% and 2% values are arbitrary, there is no perfect set of values due to the idiosyncrasies of real data. These values were tested on a subset of the total data and found to be similar to the values obtained from physical examination of the curves. When calculating the congestion at an airport, the full data set (with the top 2% added back in) was considered.

### 4.2.3 Taxi Time Simulation

In order to determine future year operating points relative to this curve, a traffic simulation capability which had been previously developed and validated at MIT [19] has been modified to use the inputs identified above to predict taxi times in the future. The simulation calculates taxi times for every flight over the course of a year in a given configuration by modeling the aircraft departure process as a queuing system. It takes the future year schedules as its main input and assumes that the scheduled departure times will be the pushback times for each flight. Taxi time, $\tau$ is related to the size of the departure queue by:

$$\tau = \tau_{unimpeded} + \alpha R(t) + W_q(t)$$

where $\tau_{unimpeded}$ is the average unimpeded time (by airline or overall), $\alpha$ is a taxiway congestion factor, $R(t)$ is the number of aircraft on ramps and taxiways at time $t$, and $W_q(t)$ is the expected waiting time at time $t$. The simulation calculates the time for three different segments of taxiing: unimpeded time, taxiway congestion time, and time in departure queue. In Figure 30 $\alpha$ represents the Ramp and Taxiway interactions, $W_q(t)$ is the time spent in the departure queue (which depends on the runway server), and $\tau_{unimpeded}$ is the base time it would take if the ramp interactions and departure queue were 0.

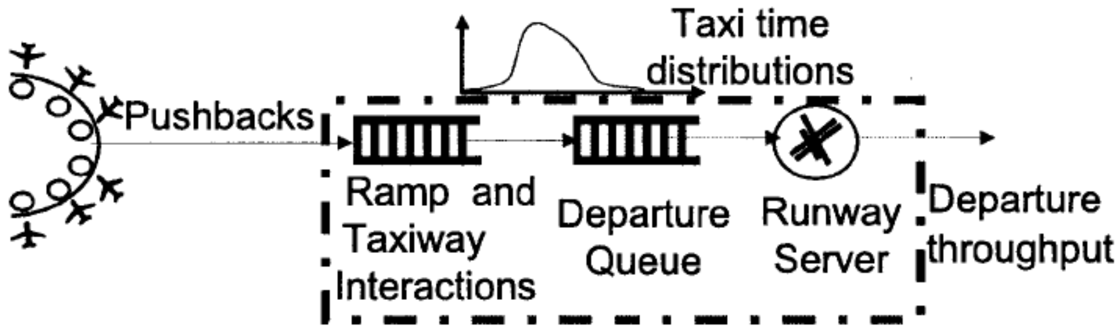



**Figure 30: Departure Process [19]**

These three segments have tunable parameters: $\tau_{unimpeded}$, $\alpha$ and capacity (which affects $W_q$). The average unimpeded time used was the average across all airlines in a specific configuration from 2010, because the physical layout of an airport can change suddenly and unexpectedly (e.g. if an airline moved terminals or left an airport). By using the overall taxi time, the robustness of the model is improved. There were no changes to the unimpeded time because of new construction because of the uncertainty in completion times and effectiveness. The taxiway congestion factor is calibrated from the present day training data by matching the amount of congestion predicted with the congestion actually seen. In [19] this factor is calibrated by matching the mean and median taxi times. Congestion matching was chosen because it more accurately predicts taxi times during congestion, which is the regime of interest.

The saturation throughput from the RF model was used to determine the service rate for the departure queue. The service rates (aircraft/minute) were calculated from the training data for different levels of arrivals to reflect the interdependence of the arrival and departure rates. The saturation throughput calculated by the RF model was an average value, so to translate that to different levels of arrivals, the difference between the average service rate and the rate implied by the saturation throughput was calculated. This difference was added to the rates for each level of arrival to determine the new service rates. The service rates were modeled as Erlang distributions, where the arrivals at the runway threshold were assumed to be random. Each runway configuration at each study airport was modeled as a single server with infinite space for the queue, and aircraft are taken first-come, first-served.

With estimates of taxi time, the evolution of $N(t)$, the number on surface, over the course of the day can be calculated. Note $N(t)$ is not the same as $R(t)$ because it includes aircraft in the departure queue at the end of the runway. Benefits of surface congestion management relative to the baseline case for future years are calculated in the same way as described previously, by taking the number of flights operating above N* and multiplying by the difference between the average taxi time at N* and the average for flights above N*.

The two main modifications to [19] that were implemented for this thesis were calibrating the taxiway congestion factor to the amount of congestion instead of the mean/median taxi times, and changing the service rates to match the values predicted by the Random Forest instead of using the values from the operational data. The second change was important because it reflects the change in airport performance due to secondary variables beyond capacity such as demand. Without it, the airport performance would be static (unless additional runways are planned, as at ORD).

In addition to the two modifications, the decision was made to model the 5 most-used configurations separately for each airport instead of choosing one ‘aggregate’ configuration. There were several reasons for this decision. The simulation is supposed to be tailored to a specific configuration with average unimpeded times and service rates. By using an aggregate configuration, one dilutes the validity of the model. In addition, if an airport has configurations that vary in performance, the benefits of SCM would be greatly affected by assuming one configuration. For example, a hypothetical airport with two configurations, one with two departure runways that is in use 55% of the time and one with one departure runway that is in use 45% of the time. If the most common configuration is used, the effective throughput of the airport is relatively high and congestion that might have been present in the configuration with one runway is eliminated.

Because the simulation is configuration-specific, realistic configuration choices are needed. The weather and configuration choices from the base year (2010) were taken as typical and used for every future year. While configuration choices and weather can change from year to year, the behavior is unpredictable and compared to any predictions one might make the base year is at least proven to be within the envelope of airport behavior.

### 4.2.4 Results, Validation, Constraints

To calculate the future benefits, demand and capacity inputs are needed. The taxi time simulation takes as a demand input a pushback schedule. This was obtained from “NextGen schedules” provided by the FAA from their SWAC model [21]. This is a NAS-wide network model which is used by the FAA to develop NextGen flight scenarios (as was used for the future schedules input discussed above), and has the ability to “trim” flights when the demand at any given node (airport) in the network causes degraded performance beyond a certain threshold level. To calculate the capacities used to trim the schedules, the average ADR in the base year was increased in the future by the percentage suggested by the MITRE FACT2 report [22]. This resulted in a discrepancy at ORD where the capacity increased by more than the average saturation throughput predicted by the RF model. This will be discussed further in section 4.4.7.

After the taxi times and saturation curves are calculated, the benefits must be summed across the 5 common configurations. Using only 5 configurations caused an inconsistency between the simulation results and the field trial results for two reasons: firstly, low-use configurations can have disproportionate benefits from metering, meaning that only scaling the top 5 configurations underestimates the true benefits. Secondly, only examining VMC conditions also underestimates the benefits because IMC conditions have more benefits / hour of time due to the reduction in capacity and congestion caused by bad weather. To account for this, the benefits for the base year were calculated from ASPM using the 5 configuration and scale up method as well as with one aggregate saturation curve that included all configurations and weather conditions. The benefits obtained using these two methods were compared to obtain a scaling factor between the scale up method and the aggregate method (more representative of actual benefits). This factor was then used to scale up the future benefits so that they were compatible with the historic benefits and the field trial.

The “unconstrained benefits” are produced by the method outlined above, but in reality there are physical constraints to the number of flights that can be held by a departure metering approach, e.g., by the number of gates or off-gate hold locations. The “practical benefits” results outlined here consider airport gates as a limiting resource. If there are too few gates, metering might need to be scaled back or conducted off gate, which is not as desirable. To calculate the gate utilization, OOOI times from ASPM were used. The result is that the approximate number of aircraft on the ground (assumed to be at a gate) can be calculated throughout the day. This count is calculated by adding one when an aircraft arrives at a gate (IN time) and subtracting one when an aircraft departs (OUT time). The count is calculated at each minute from midnight to midnight of one day and is airline-specific. Finally, because the count starts at midnight there are an unknown number of aircraft already on the ground. This results in a count that can be negative at times. To normalize for this, the absolute value of the minimum value (largest negative number) is added to the entire count for that airline. For example, if an airline has a point in time where the count is -8, 8 would be added to every value in the count so that the new minimum value is 0. This approach assumes that each airline has 0 aircraft at a gate at one point during the day. This not completely accurate, but the induced error is small. The capacity of the airport to conduct on-gate holds can be estimated by taking the difference between the number of gates in use and the total number of gates at the airport. This method makes several simplifying assumptions: It neglects gate ownership issues (in the US, gates are ‘owned’ by a specific airline and are not a shared resource), the size of gates and their ability to handle different types of aircraft and whether or not an aircraft was moved off gate after arrival,. It also does not explicitly show space available for off-gate holds. Off-gate holding space is very hard to quantify without interviews with staff at specific airports, but examination of LGA maps (the most constrained airport) identified several possible locations. We therefore assumed that off-gate holds could be used at all airports. With this assumption, our cutoff for on-gate metering was the total number of gates at an airport. We assumed that the utilization could temporarily hit that peak, with the excess demand held off gate, but that if the utilization was significantly over the number of gates, metering could not take place.

The gate utilization was calculated for each airport and year in the study and compared to the number of gates at the airport (or planned to be constructed). If the analysis showed that there would not be enough gates to accommodate metering, then the benefits were restricted to the last year in which the gates could accommodate metering. Figure 31 shows results using future schedules from DFW and JFK in 2010. The average number of aircraft at a gate is shown on they axis, with the X-axis showing the time of day. Both airports have approximately 150 gates but face different future demands. While DFW is forecast to have little growth in demand for gates in the future, JFK will, according to these schedules, face increasing competition for gates even without the implementation of SCM.

Several other airports will be shown to exceed their gate capacity in future years. This illustrates a fundamental problem with the generation of future schedules: the only constraining capacity is the runway capacity when there are in fact several others that can restrict an airport, such as gate capacity, security, and noise abatement. Because these factors are not considered, the use of these schedules can lead to overestimates of benefits because demand levels are higher than realistic levels at several airports. While we attempt to correct for this by restricting growth of benefits when gate constraints are met, the preferred method would be to regenerate the schedules with additional constraints. Unfortunately, this was not feasible for this thesis. Because the methodology behind the generation of the future schedules is unknown, the impact on the results is hard to quantify. Nevertheless, results using a new set of schedules would not be expected to be substantially different from the 'Practical' set of results because they reflect the benefits at the maximum sustainable demand given the shape of the schedule.

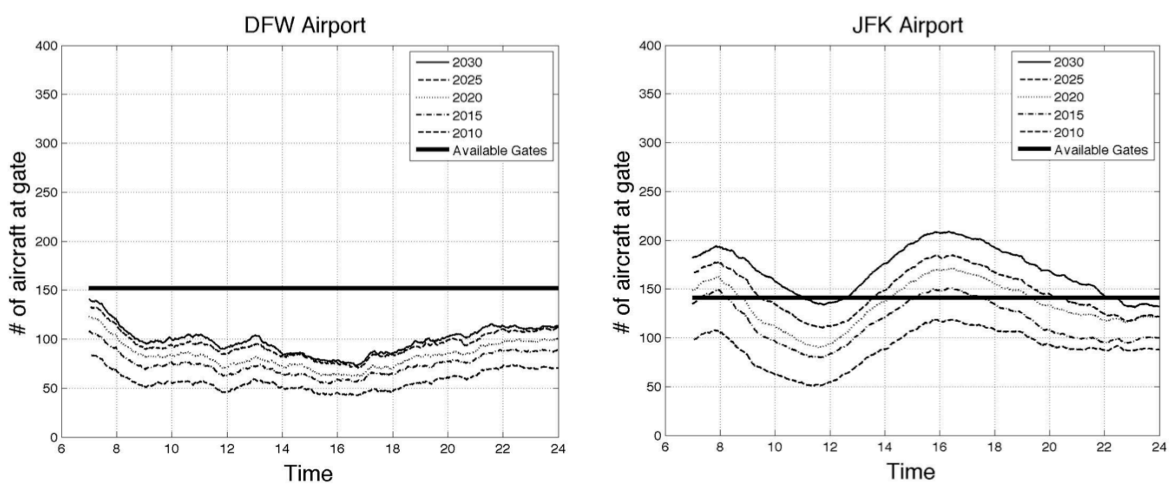



**Figure 31: DFW and JFK Gate Utilization (Future Schedules)**

## 4.3 AIRPORT SCOPE

8 airports have been examined in detail: ATL, BOS, DFW, IAD, JFK, LGA, ORD, and PHL. JFK and BOS were chosen because of their recent and ongoing field trials of SCM. The other airports were chosen to represent different types of airports. LGA is small and space constrained, PHL is larger but space constrained, DFW is large with relatively low demand, ORD and ATL are large with high demand, and IAD is a medium sized airport. Results for each airport are included in the next section.

## 4.4 RESULTS

Results are given for the 8 study airports in terms of hours of taxi time reduction. In the discussion section, these time savings will be converted into fuel and monetary savings. The predicted benefits are shown in panel b for each airport. The historical benefits (Benefits that could have been realized if metering was in place) between 2000 and 2010 were calculated with ASPM data and are displayed for comparison. These historical benefits estimates generally validate the methodology because the 2010 (Actual) and 2010 (Simulated) points are close. When field demonstration data is available (i.e. for JFK and BOS), that too is used for comparison. The top-left chart for each airport shows different measures of demand and capacity between 2000 and 2030 that helps to interpret the unconstrained benefits results. The historical capacity is the average declared departure capacity from ASPM, while the saturation throughput is calculated from the average of the throughputs for the saturation curves for the 5 most used VMC configurations. The bottom left chart for each airport shows the average gate utilization during the study years, as well as the current number of gates at the airport. This is used to identify disparities between forecast demand and forecast gate availability. If there are significant gate constraints, then the benefits are capped and the bottom right chart shows the resulting practical benefits.

### 4.4.1 ATL - Atlanta

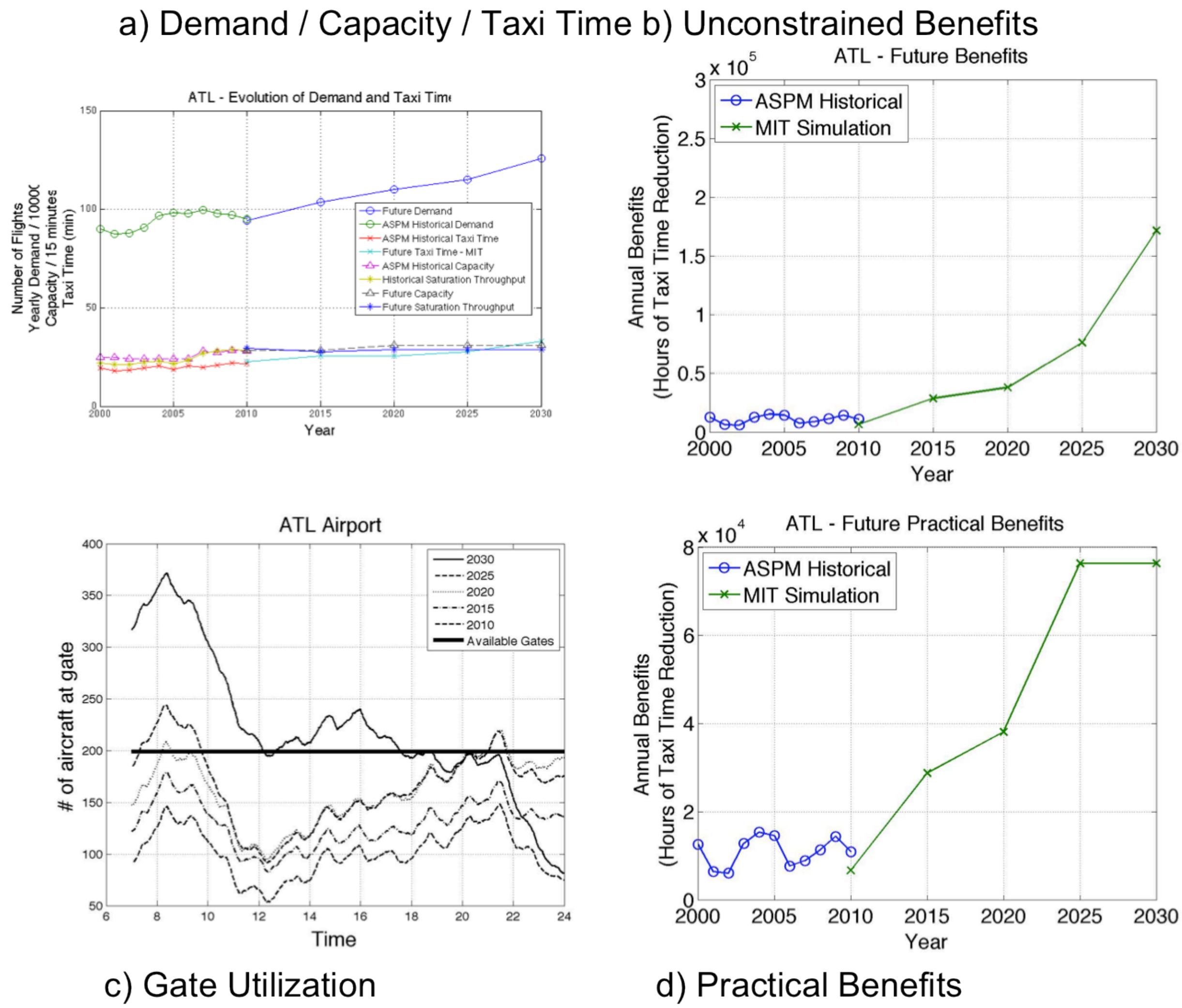



**Figure 32: ATL Departure Metering Benefits Results**

Demand at ATL is predicted to increase by about 30% from 2010 to 2030, while the predicted capacity does not increase significantly. The historical benefits show a large increase in 2003-4 as a result of the increasing demand, before dropping off as the airport performance improved from 2006-8 (see the saturation throughput / capacity) with largely steady demand. As a result of the increased demand in the forecast, the airport shows steadily increasing unconstrained benefits from departure metering through 2030 The gate utilization results show that there is significantly higher need for holds than there are gates available in 2030, so the practical benefits are capped at 2025 levels. It is assumed that the relatively short time periods in 2025 where gate utilization exceeds gate capacity can still be metered by holding aircraft in off-gate locations. The unusual gate utilization in 2030 will be explained in the discussion section. The gate utilization for all airports shows a steady scaling up of the current pattern with increasing demand. Analyzing the shape of the curve is beyond the scope of this thesis, but instead of simply scaling up a future utilization curve might be flatter over the course of the day if an airport is near capacity (akin to current day LGA).

### 4.4.2 BOS – Boston Logan

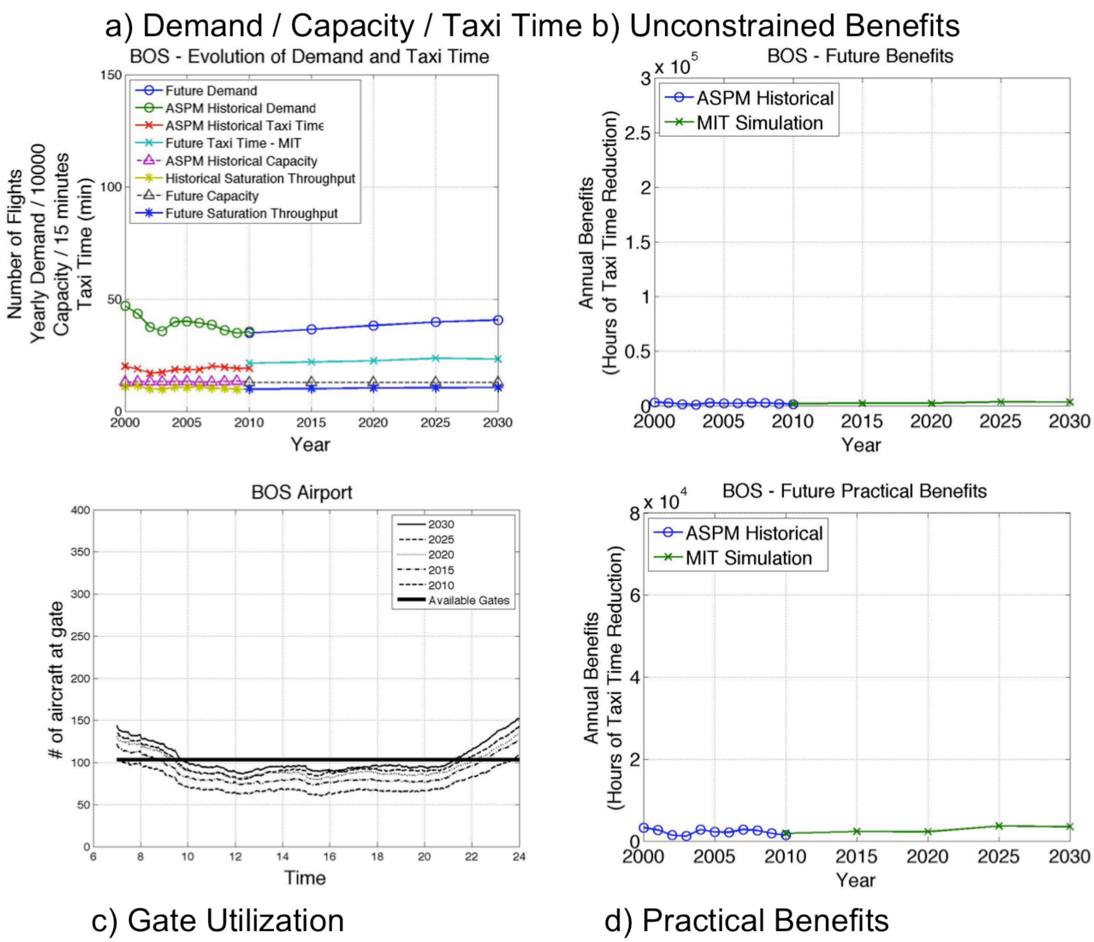



**Figure 33: BOS Departure Metering Benefits Results**

Expected demand growth at BOS is about 20% from 2010 to 2030 (although demand in 2030 is not expected to be any larger than the airport handled in 2000), while capacity is not expected to increase significantly. The historical benefits roughly follow the historical demand at BOS, but are very low relative to the other study airports. Because the demand does not reach the levels seen in 2000, it is reasonable that the benefits also stay within the range already seen. Gate utilization is not expected to be a constraining factor at this airport because it remains under its capacity for the duration of the day and therefore the practical benefits are expected to be similar to the unconstrained. The extrapolation of the results of the MIT field trial previously discussed in Chapter 2 is shown in Figure 34. The results from the simulation for the base year of 2010 have been added to the chart and compare favorably to the configuration specific benefits estimated from both the field trial and ASPM saturation curve for 2010, providing some validation of the operational realism afforded by the approach.

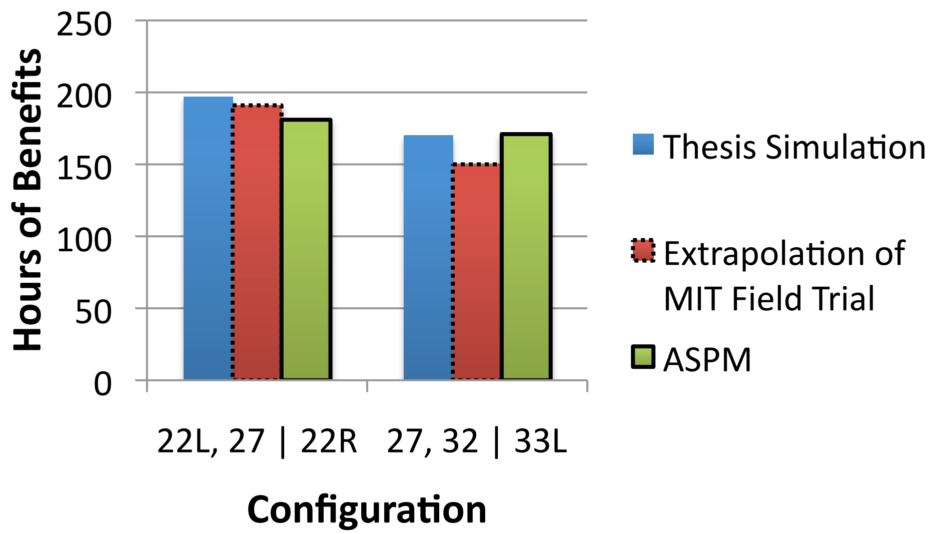



**Figure 34: Boston Field Demonstration Validation**

### 4.4.3 DFW – Dallas / Fort Worth

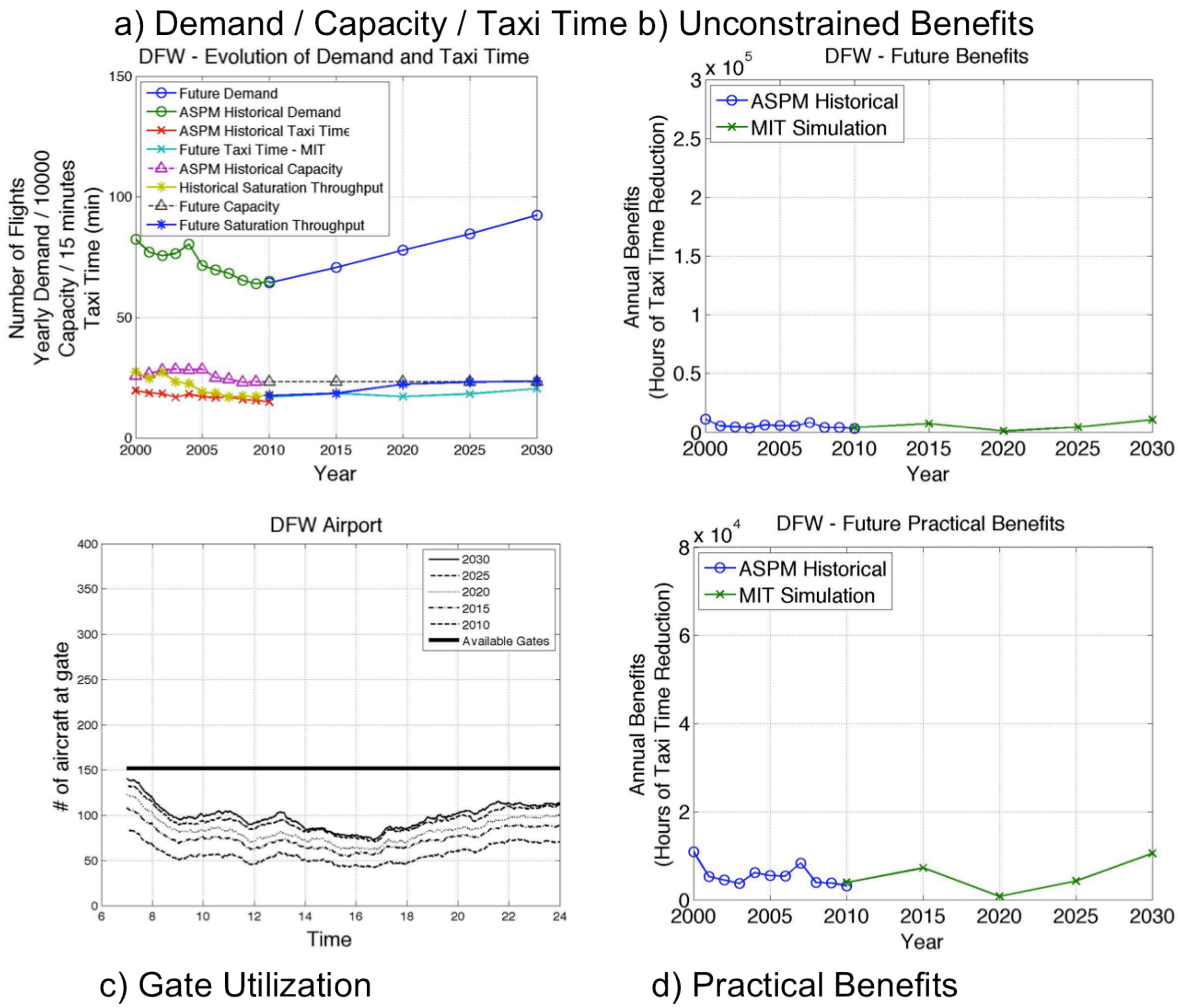



**Figure 35: DFW Departure Metering Benefits Results**

DFW airport is forecast to reverse its recent decline in traffic, and there is no expected increase in capacity. The decrease in the future benefits in 2020 is due to the simulation-predicted increase in performance (and corresponding decrease in benefits) between 2015 and 2020, which can be seen in panel b. While this may seem like a sudden change, the historical data shows substantial volatility. This is a reversion to the performance in the early 2000's, when there was a similar demand level. This effect is elaborated on in section 4.6.1. There should be sufficient gate space without any expansion in terminal facilities (and there is room for terminal expansion at the airport if necessary), resulting in practical benefits that are equal to the unconstrained benefits.

### 4.4.4 IAD - Washington Dulles

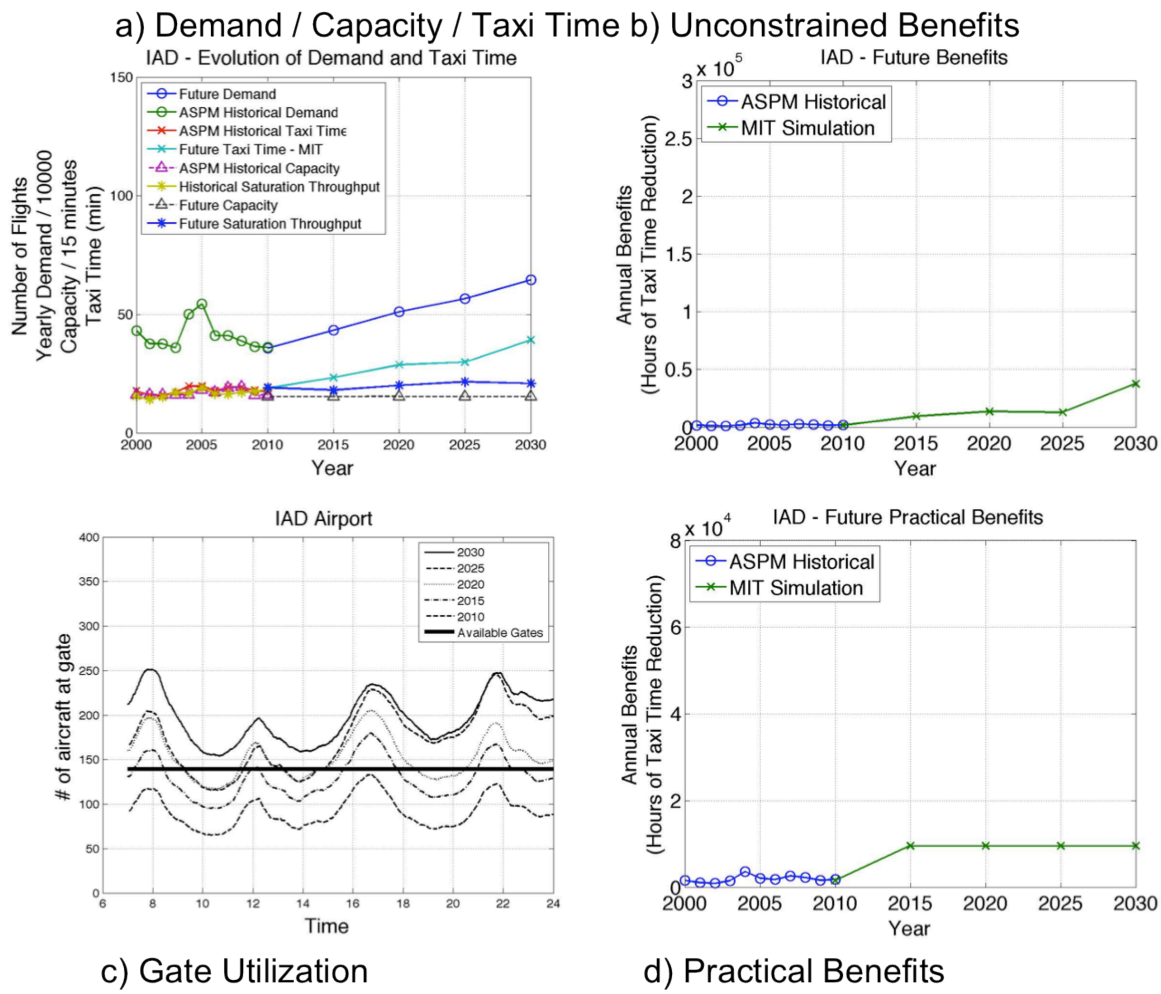



**Figure 36: IAD Departure Metering Benefits Results**

Dulles airport shows steadily increasing traffic from 2010 through 2030. There are several discrepancies in the results that merit explanation. The first is the spike in traffic in 2004 and 2005 that did not produce the same magnitude increase in taxi time and benefits relative to other airports. The sharp rise in traffic was due to the rise and fall of Independence Air, which was based at Dulles and operated from June 2004 to January 2006. The gate usage chart in Figure 38 from 2004 suggests that Independence tried to avoid the departure banks of United that would minimize the added congestion and benefits. Compared to the 2010 chart in Figure 38 (uses ASPM data as opposed to Figure 36 which uses the future schedule), there are secondary departure banks around 0700, 1000 and 1400 in between the main United pushes. This would mean that congestion is lessened because the increased demand happened at off-peak times.

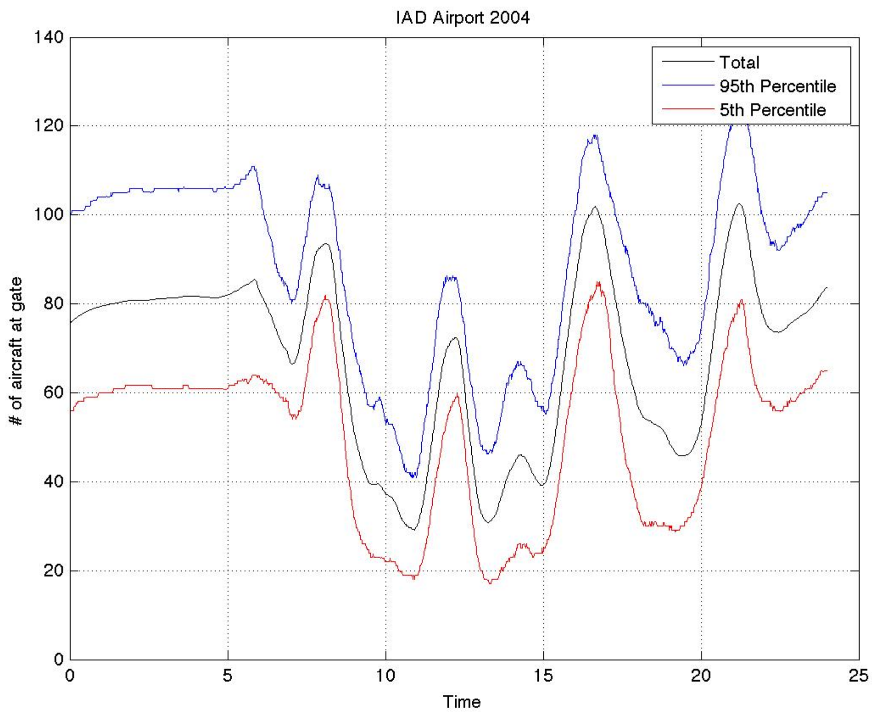



**Figure 37: 2004 Gate Usage at IAD**

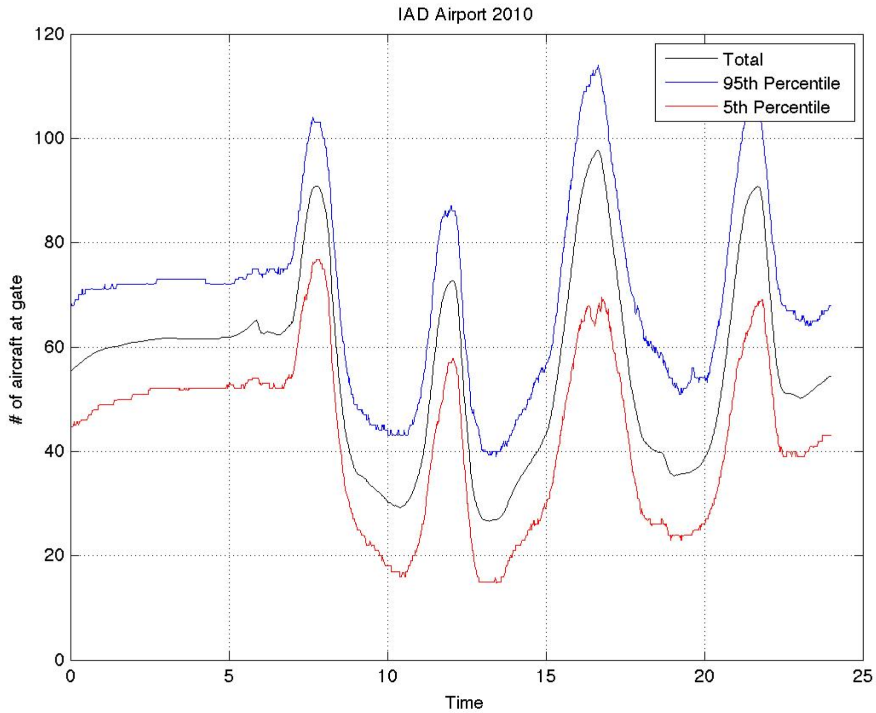



**Figure 38: 2010 ASPM IAD Gate Usage**

Traffic in 2020 is predicted to be at 2005 levels and the predicted benefits are much higher most likely because the growth is spread evenly across all carriers and the timing of the additional flights is not focused in off-peak times as it was in 2004. Increasing traffic during the busy departure pushes would greatly increase congestion and benefits, much more so than the addition of Independence Air in 2004-5. A second cause for the future benefits being higher than the present day is shown in Figure 39. In the ASPM individual flights database, there are no days with more than 500 departures while the FAA schedule has multiple such days. The ASPM historical demand shown in Figure 36a matches with the future demand because it is an aggregate count that includes flights such as military and GA that are not necessarily counted in the individual flights database. This makes a difference because the individual flights database is used to calculate the saturation curve and taxi time simulation. The effect of the difference between the aggregate and individual databases is negligible at the other airports, but IAD has a 17% difference in the number of flights (BOS is second at 6%). One solution would be to further trim the demand at IAD but that was not possible for this thesis. In terms of gate utilization, Figure 36 shows that there are significant gate conflicts from 2020 onwards, especially with the ‘banked’ behavior due to the United hub. Benefits are thus capped at 2015 levels. The IAD saturation curves from 2005 and 2006 (Figure 40 and Figure 41) support the hypothesis that IAD has not reached its capacity, and are very different even though they are only one year apart due to a 20% drop in VMC traffic. The difference between the curves shows that IAD appears to not have reached its maximum capacity even at 2005 traffic levels. In 2005, the calculated N* is 30 even though the throughput continues to increase after that point. This is due to the method of calculating N* which discards the top 2.5% of flights as unreliable. In this case, however, there appears to be a definite trend showing possible higher performance. The current methodology does not capture the possible higher performance at high demand levels, leading to increased estimates of congestion in the future. Another problem with IAD in particular is the reporting of its configurations. The saturation curves are for 1L, 1R | 30 which means that arrivals are on 1L and 1R with departures on 30. For a typical airport with one departure runway (LGA, JFK), the saturation throughput is around 10 / 15 minutes. The saturation throughput here is between 18 and 25 / 15 minutes, suggesting that the airport is using other runways as departure runways and not accurately reporting it. This decreases the accuracy of both the RF model and the traffic simulation.

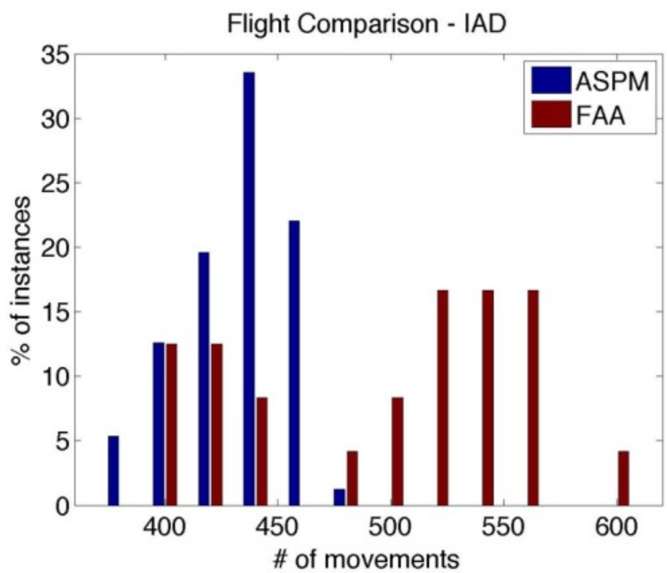



**Figure 39: Average number of departures / day in ASPM and FAA schedule (2010)**

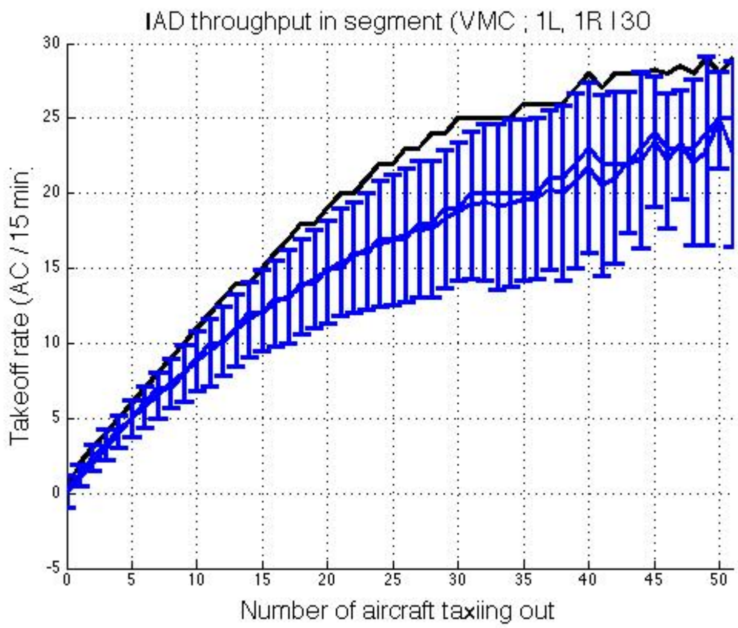



**Figure 40: IAD 2005 Saturation Curve**

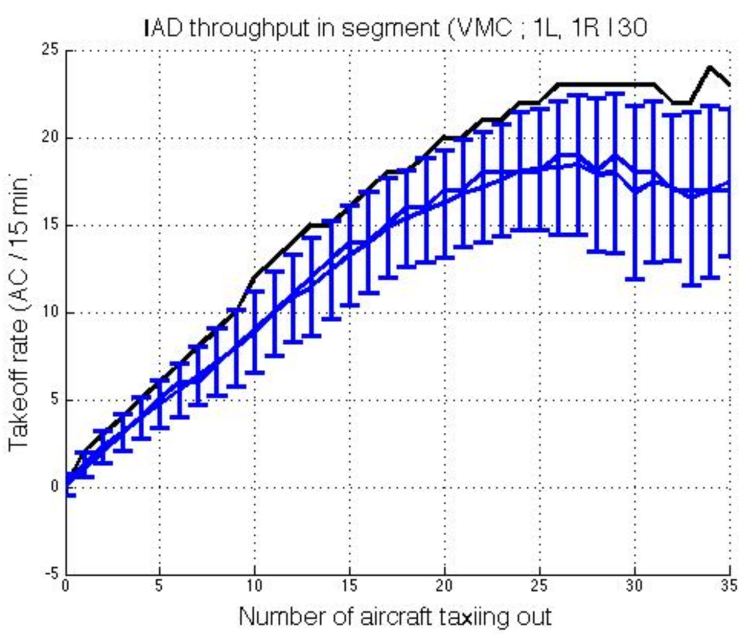



**Figure 41: IAD 2006 Saturation Curve**

### 4.4.5 JFK – New York JFK

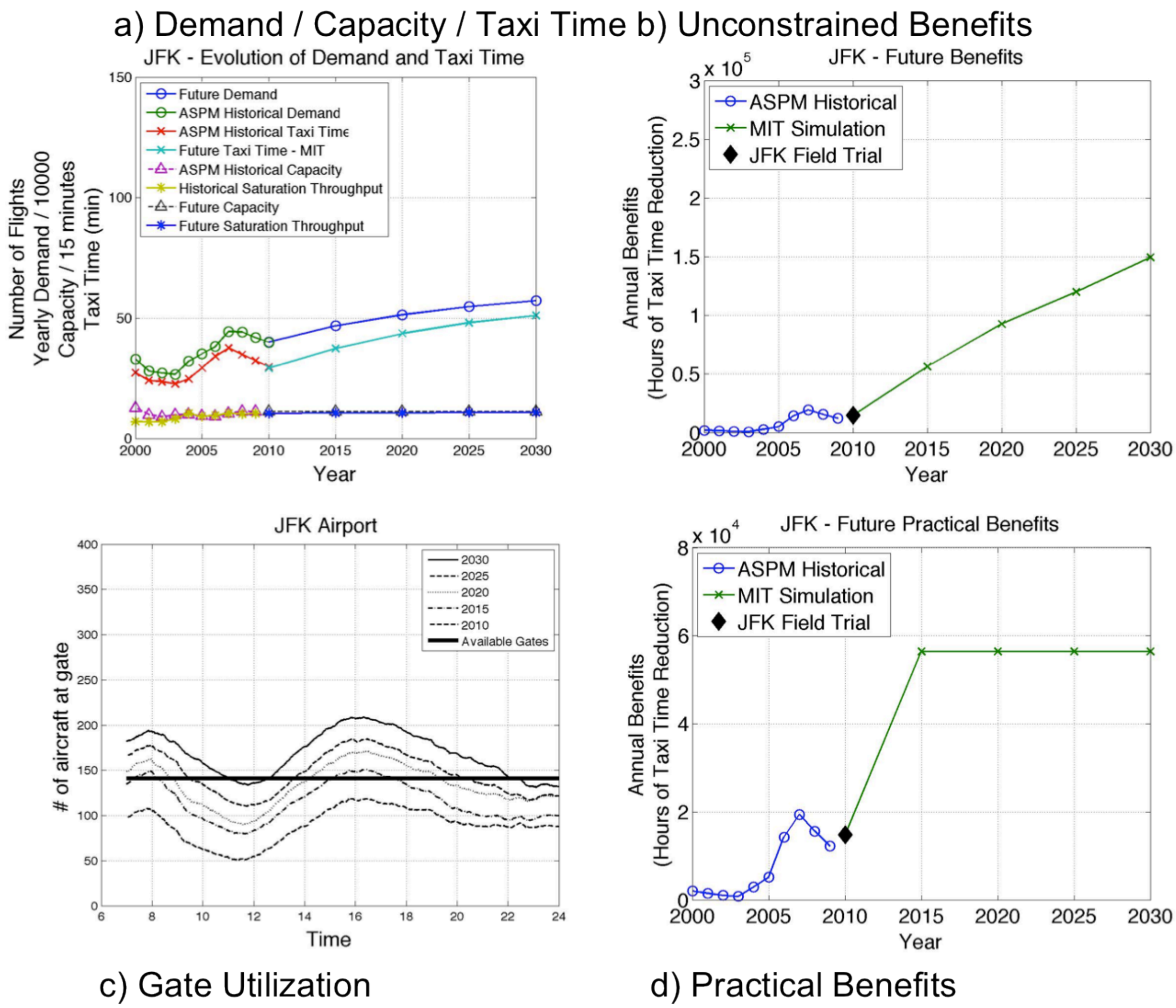



**Figure 42: JFK Departure Metering Benefits Results**

JFK shows steadily increasing benefits as the demand increases. The demand is substantially above the peak historic traffic level seen at JFK in 2007 even though the capacity is not forecasted to grow and taxi times and delays are already high. The saturation throughput is forecast to grow slightly as a response to the increasing demand. The growth in demand coupled with a lack of growth in capacity results in large benefits into the future, but the demand for metering will exceed the available gate space soon after 2015. This indicates that either there needs to be new terminal construction (which is not planned) or that the demand is too high. Given this constraint, the practical benefits from metering are capped at 2015 levels. There is no data point for the historical benefits in 2010 because of the field demonstrations being conducted at the airport by PASSUR at that time. However, these trials allow further validation of our approach. The comparison between the benefits from the field results (the black diamond in panel d) and ASPM data from Chapter 3 is reproduced here with the addition of the simulated benefits.

### 4.4.6 LGA – New York LaGuardia

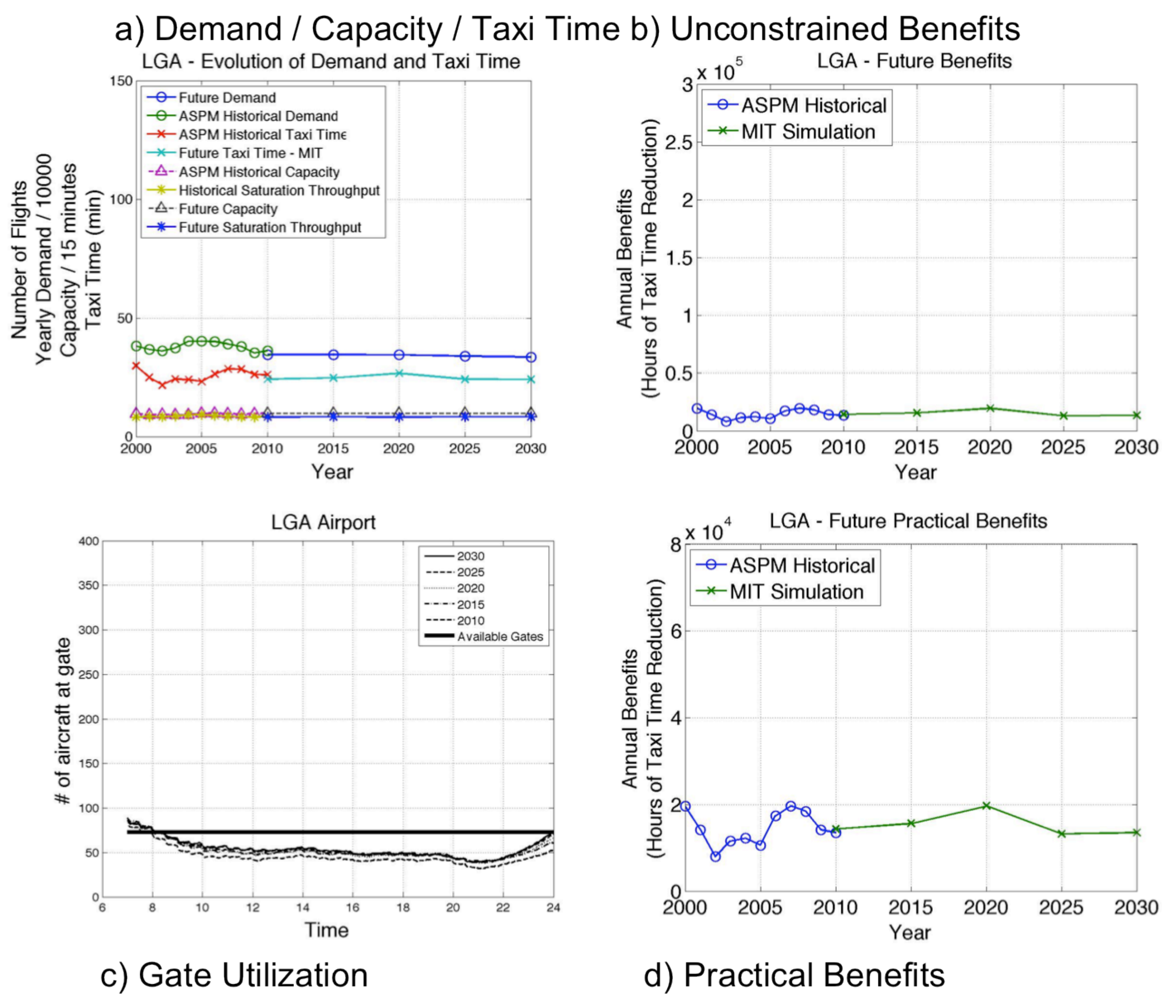
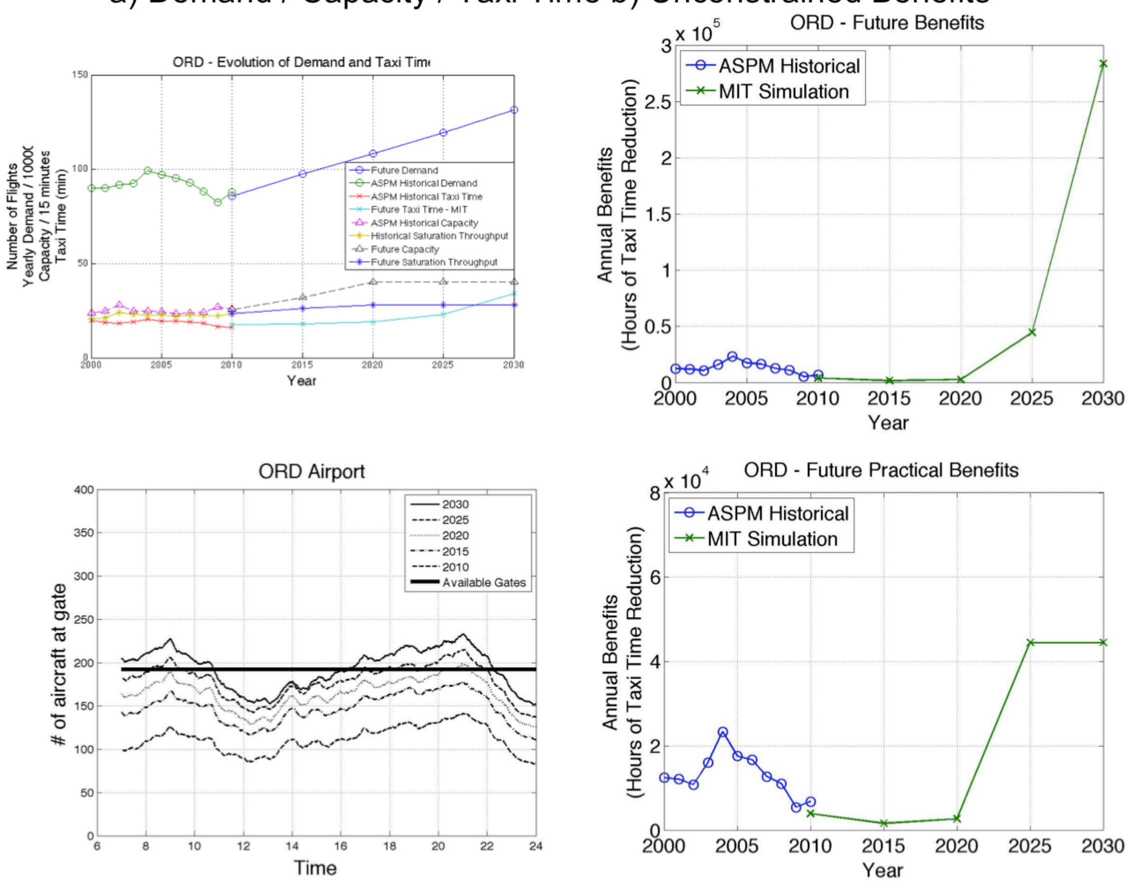



**Figure 43: LGA Departure Metering Benefits Results**

LaGuardia airport is forecast to have no growth in the study time period. This is expected given LGA's slot-controlled status. As a result, the benefits do not vary much over the course of the study. The spike in 2020 is likely due to small variations in the predicted performance of the airport (the average taxi time is slightly higher) and does not have a major impact on overall benefits. There are a significant number of open gates during the day in all years, allowing for the full benefits of metering to be achieved. This figure is roughly similar to Figure 31, which serves as a reality check.

### 4.4.7 ORD – Chicago O'Hare

a) Demand / Capacity / Taxi Time b) Unconstrained Benefits




c) Gate Utilization d) Practical Benefits

**Figure 44: ORD Departure Metering Benefits Results**

O'Hare airport has relatively low benefits through 2020 because of the runway capacity expansion project scheduled to finish by 2020. This is the only airport in the study undergoing physical expansion. The expansion introduces much uncertainty into the calculation of benefits because the capacity of the airport determines the demand (through trimming) but the saturation throughput (performance) determines the benefits. The simulation implies high levels of benefits in 2030 because it predicts that the saturation throughput of ORD will not be as great as the predicted capacity. This prediction is based on the past performance of ORD as well as the performance of DFW, an airport whose current configuration is similar to ORD's future configuration. The demand will thus exceed the capacity at ORD in the model, causing a large amount of congestion. The future ORD will be comparable to the current DFW layout. The peak DFW throughput, which occurred in 2002, was 28 departures / 15 minutes, which is around the simulation level. If this is the true performance of ORD in the future, the demand will likely be forced lower than forecast to keep delays low, which would in turn drive the benefits lower. If, instead, the airport performs at the level predicted by the capacity, the demand would remain the same but the benefits would again be lower because there would be fewer delays. Because in either case the benefits would be lower than the simulation predicts, the practical benefits for 2030 were capped at the 2025 level. Demand for metering can be satisfied with current terminal infrastructure until 2025. However, ORD has several locations identified for future terminal expansion. Therefore, it is assumed that there will be sufficient gate space to perform metering in 2030 with the given demand.

### 4.4.8 PHL – Philadelphia Intl

### 4.4.9

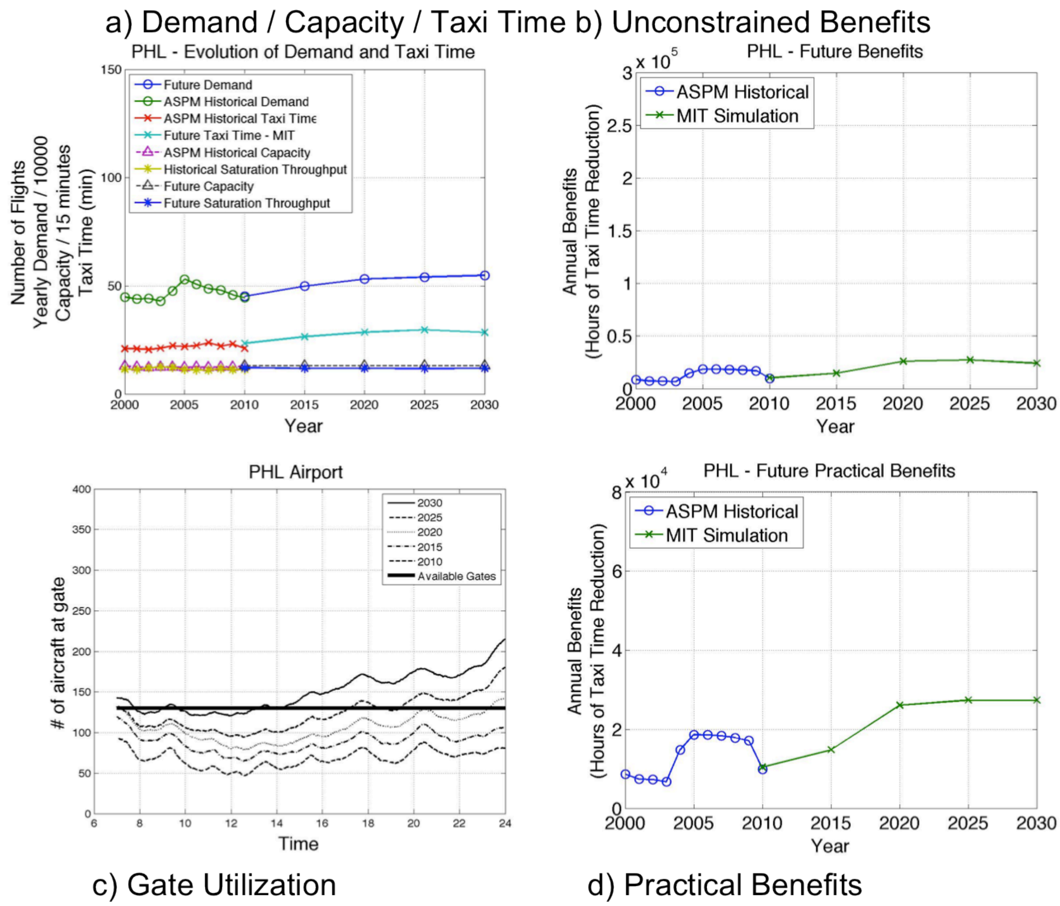



**Figure 45: PHL Departure Metering Benefits Results**

Philadelphia airport shows a medium level of benefits comparable to other airports in the study, with the future demand in 2030 slightly exceeding the peak seen in 2005. The main anomaly is the unusual behavior of the gate utilization curves for 2020, 2025 and 2030, which all end the day with many more planes than they started with. Panel c shows that there will be insufficient gate space in 2030 to accommodate the demand. Therefore, benefits are capped at 2025 levels. This causes the benefits in 2030 to rise because the airport performance was predicted to incrementally improve from 2025 to 2030, lowering the benefits relative to 2025.

When the demand is capped at 2025 levels, this improvement in performance is nullified and the benefits increase to 2025 levels.

## 4.5 AGGREGATE BENEFITS

### 4.5.1 Aggregate Departure Metering Benefits

The airports with major contributions to the unconstrained benefits are ORD, ATL, and JFK, as shown in Figure 46. The other four airports are all at about the same lower level of benefits. When the practical benefits are examined which account for gate constraints, JFK and ATL approach the lower group, as shown in Figure 47. ORD has reduced benefits because of the uncertainty in the airport expansion project. The practical benefits show that all airports are in the 0-50,000 hours of taxi time reduction range. This is a realistic value. Taking JFK 2015 as an example: 50,000 hours of taxi time reduction means that there were about 50,000 hours of gate hold in a year. That equates to 137 hours a day.

At an average of 640 departures a day, on average a flight will be held 13 minutes. Because metering will most likely be needed (and used) only at certain times of day, average holds could be between 20 and 30 minutes. Departure metering does not reduce delays, but only transfers them to the gate. Therefore, holds of between 20 and 30 minutes implies delays of at least that long. The amount of delay implied by the 2030 JFK unconstrained benefits (approximately 3 times 2015 or 1.5 hours), would likely lead to a reduction in demand until reasonable delay levels were reached.

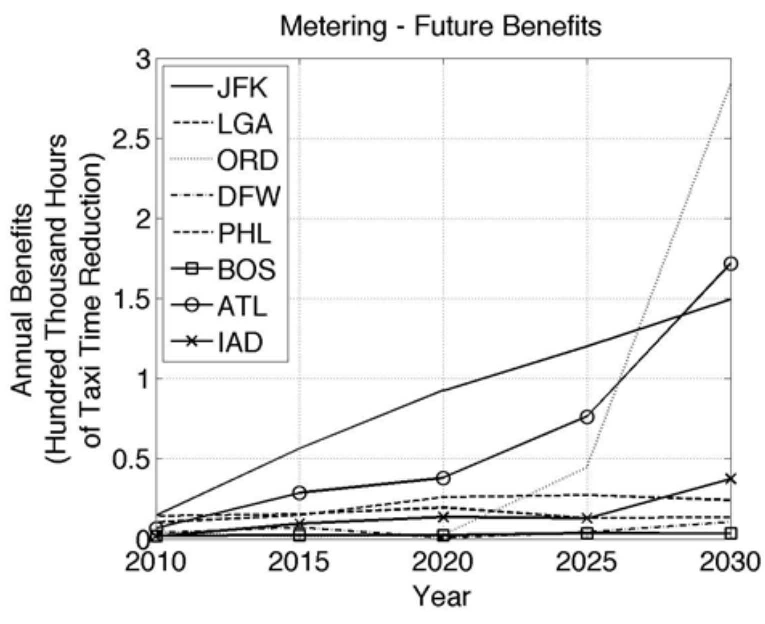



**Figure 46: Aggregate Unconstrained Results**

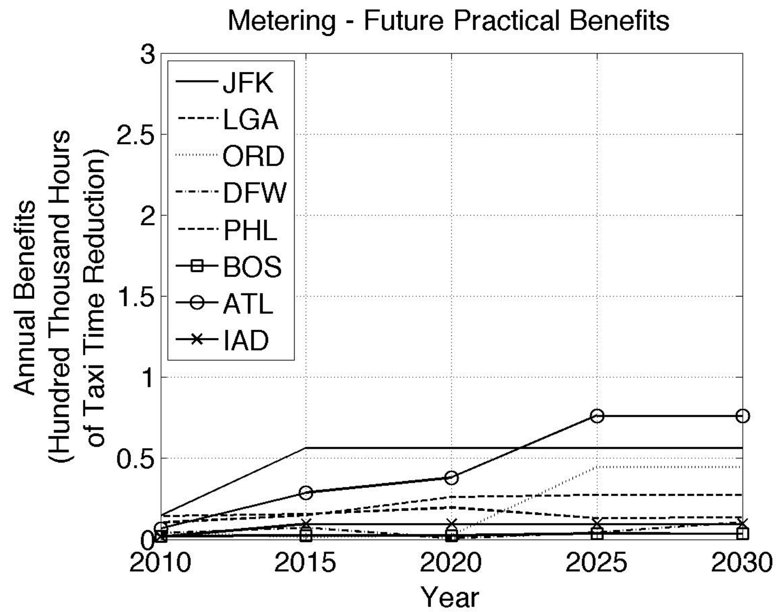



**Figure 47: Aggregate Practical Results**

### 4.5.2 Monetization of Departure Metering Benefits

Assuming a fuel price from 2010 to 2030 of $2.43 in FY2011 dollars[23], the unconstrained benefits from fuel savings can be calculated to be $3.6 billion dollars cumulative from 2010 to 2030 across the 8 airports studied, as shown in Table 3 and Table 4. Summing the practical benefits results in $2.3 billion. These estimates assume 3.1 kg / gallon of jet fuel and airport-specific fuel burn rates (using ICAO taxi fuel rates) which account for the fleet mixes at each as shown in Table 2 below. Taking the average taxi times from the simulation and multiplying by the total number of flights, we can find the total estimated time spent in taxi as well as the corresponding fuel burn. We calculated that for the 8 airports cumulatively from 2010 to 2030 in the Unconstrained case there would be 6 billion gallons of fuel burned, or $14.4 billion, making the savings from SCM almost 26% of the total fuel cost in taxi. This is higher than the estimate presented in Chapter 3 because these eight airports are forecast to operate in congestion much more frequently in the future, requiring SCM to operate for longer periods of time. In terms of total fuel burn in all stages of flight, scaling data for 2010 from the BTS [18] by the future demand levels, we can calculate that the total fuel burn at these 8 airports will be 100.5 billion gallons, with a corresponding cost of $244 billion. The benefits are then 1.5% of the total fuel burn. For the practical case, the results are 5.3 billion gallons of fuel, $12.8 billion, and 18% of total fuel cost for taxiing only and 97 billion gallons of fuel, $235.8 billion, and 1.0% of the total fuel burn. Again, care should be taken with the estimate of the percentage of total fuel burn because only domestic carriers are included, and the true percentage is probably lower. The percentages vary substantially by airport because of the nature of SCM. For airports such as BOS with little potential for benefits, most of the taxi time will not be during congestion and therefore the percentage of taxi (and total) fuel saved is low.

**Table 2: Airport-Specific Fuel Flow Rates**

| Airport | Fuel Burn Rate (kg/sec) |
|---|---|
| ATL | 0.2155 |
| BOS | 0.1892 |
| DFW | 0.2214 |
| IAD | 0.1729 |
| JFK | 0.3096 |
| LGA | 0.1707 |
| ORD | 0.2099 |
| PHL | 0.1733 |

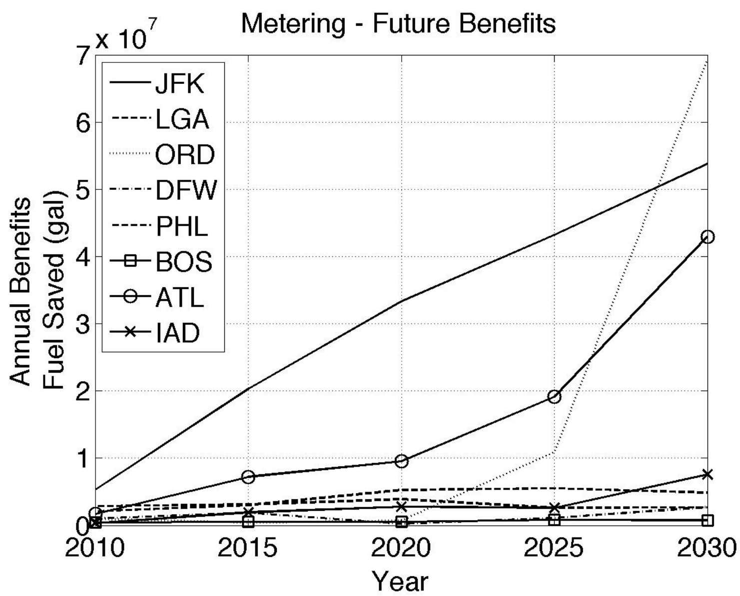



**Figure 48: Unconstrained Fuel Benefits from Metering**

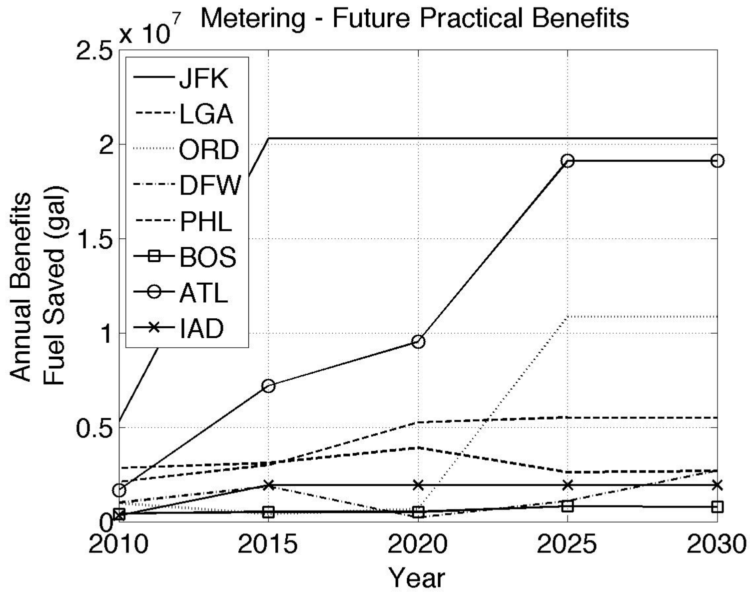



**Figure 49: Practical Fuel Benefits from Metering**

**Table 3: 2010-2030 Cumulative Benefits by Airport, Unconstrained**

| Airport | Unconstrained | | | | |
|---|---|---|---|---|---|
| | Thousand Hours Taxi Time Reduction | Million Gallons | $ Millions | Savings as % of taxi-out fuel cost | Savings as % of total fuel cost |
| ATL | 1251 | 313 | 761 | 26% | 1.5% |
| BOS | 59 | 13 | 31 | 4% | 0.2% |
| DFW | 105 | 27 | 66 | 4% | 0.2% |
| IAD | 299 | 60 | 146 | 12% | 0.6% |
| JFK | 1839 | 661 | 1606 | 47% | 6.8% |
| LGA | 326 | 65 | 157 | 22% | 1.0% |
| ORD | 1108 | 270 | 656 | 26% | 1.3% |
| PHL | 446 | 90 | 218 | 20% | 0.9% |
| TOTALS | 5,432 | 1,498 | 3,641 | | |

**Table 4: 2010-2030 Cumulative Benefits by Airport, Practical**

| Airport | Practical | | | | |
|---|---|---|---|---|---|
| | Thousand Hours Taxi Time Reduction | Million Gallons | $ Millions | Savings as % of taxi-out fuel cost | Savings as % of total fuel cost |
| ATL | 965 | 242 | 587 | 21% | 1.2% |
| BOS | 59 | 13 | 31 | 4% | 0.2% |
| DFW | 105 | 27 | 66 | 4% | 0.2% |
| IAD | 177 | 36 | 86 | 11% | 0.6% |
| JFK | 1060 | 381 | 926 | 35% | 1.9% |
| LGA | 326 | 65 | 157 | 22% | 1.2% |
| ORD | 390 | 95 | 231 | 10% | 0.5% |
| PHL | 455 | 92 | 223 | 20% | 1.1% |
| TOTALS | 3,537 | 949 | 2,307 | | |

## 4.6 ANOMALIES AND DISCREPANCIES

### 4.6.1 Uncertainty in Runway Capacity / Performance

Impacts: ORD, DFW

Both ORD and DFW have varying estimates on the future capacity and performance of the airport. In the case of ORD, new construction will add runways but the usage and performance of the new configurations is unknown. DFW has no new construction, but history has shown large variations in performance with changes in demand. It is not certain in either case what the future performance will be, but it will have a major impact on the benefits. DFW will be less impacted because the demand at the airport is not forecast to reach even the conservative estimate of capacity, keeping the benefits levels low. However, the volatility in the results is visible in Figure 35. ORD, on the other hand, would likely not be able to sustain operations at the demand level of 2030 because the 300,000 hours of gate hold in 2030 translates to about 40 minutes for each flight. This is not an acceptable level of delay (would likely translate to average taxi times over 1 hour without metering).

DFW shows the most variation in its historical performance out of the eight study airports. There are several hypotheses on why this has happened. The first is that the large decrease in demand allowed throughput to drop without increasing delay. Without the pressure of high demand levels, the airport did not need to perform at high levels and so did not.

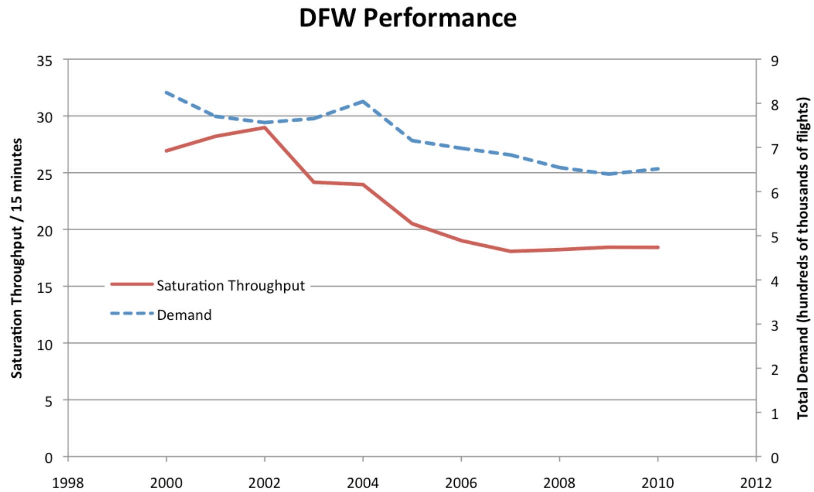



**Figure 50: DFW Demand and Performance**

Figure 50 shows the saturation throughput of the main configuration (South Flow, 13R, 17C, 17L, 18R | 13L, 17R, 18L) from 2000 to 2010 as well as the total number of operations at DFW in each year. The saturation throughput appears to be correlated to the total demand most strongly from 2004 onwards. However, the period between 2000 and 2004 cannot be readily explained by the demand because it decreases from 2000 to 2002 while the throughput increases and then increases from 2002 to 2004 while the throughput decreases.

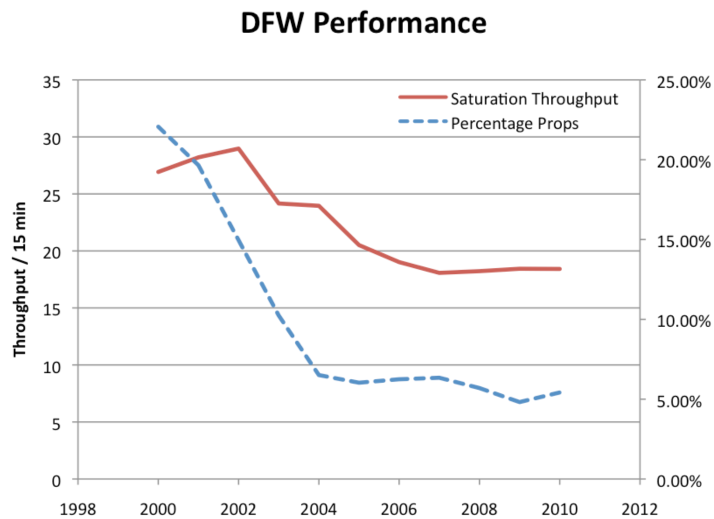



**Figure 51: DFW Proportion of Prop Operations**

Figure 51 shows the same saturation throughput, but compared to the percentage of operations that are propeller-powered aircraft. This is a relevant comparison to make because Runway 13L is constrained to prop-only operations for noise abatement. When prop operations made up a substantial proportion of operations (until 2004), the throughput was higher, between 24 and 29 departures / 15 minutes. The typical performance of a single departure runway is around 10 departures / 15 minutes. Because the south flow has three departure runways, one would expect the maximum throughput to be around 30, depending on runway crossings, interactions with landings, and symmetrical use of all 3 runways. However, when prop operations sharply fell off, the effect was to lose a runway because jet operations could not use 13L.

The behavior between 2000 and 2004 is still not fully explained, but it demonstrates the complexities and unpredictability of airport behavior. In the future we assume the same fleet mix as present day because it is the best guess, but at airports like DFW there could be major changes if the demand approaches the capacity. Airlines could see a runway lying essentially unused, and convert many of their regional flights back to turboprop aircraft to take advantage of the unused capacity, bumping the throughput back to 2000 levels. Alternatively, airport officials could decide to lift or partially lift the props-only requirement. These choices would have major impacts on the benefits (and, in effect, are implied by the RF model when it predicts that the throughput will rise in 2020), but cannot be predicted.

The prop-only restriction and change in fleet mix was only found after noticing the large change in throughput from 2000 to 2010 and investigating further, but it can easily be imagined that there are other such operational changes that have smaller effects and are thus neglected. These add to the uncertainty of trying to predict future behavior but cannot be discovered without intense inspection of each airport studied, which, given the medium and low fidelity methods for estimating future benefits, is not desirable or efficient. It is important to consider the large amount of uncertainty inherent in any future prediction, and especially in this particular case. A section of Chapter 6 will identify and discuss the main sources of uncertainty in this thesis.

### 4.6.2 Gate Utilization

Impacts: LGA, BOS, ATL

Both LGA and BOS show an anomalous hump in the overnight gate utilization that is much higher than is seen in the ASPM 2010 data. This is because the FAA future schedules have a different distribution of arrivals and departures than the actual distribution in 2010, shown in Figure 52. Because the methodology behind the future schedules is unknown, it is assumed that it is more likely that the gate utilization will resemble the current day pattern. While this difference in distributions could cause a change in the benefits level, it is outside the scope of this project to calculate a new schedule with more realistic distributions. The 2030 data for Atlanta does not follow the trend for the previous years mainly because of a substantial imbalance that appears between the departures and arrivals. Whereas in 2025 there are predicted to be 20,000 more arrivals than departures, in 2030 at ATL there are predicted to be 60,000 more departures than arrivals. Because the cause of this discrepancy lies in the generation of the future schedules, there is little that can be done. The true gate usage at ATL will likely follow the current day pattern.

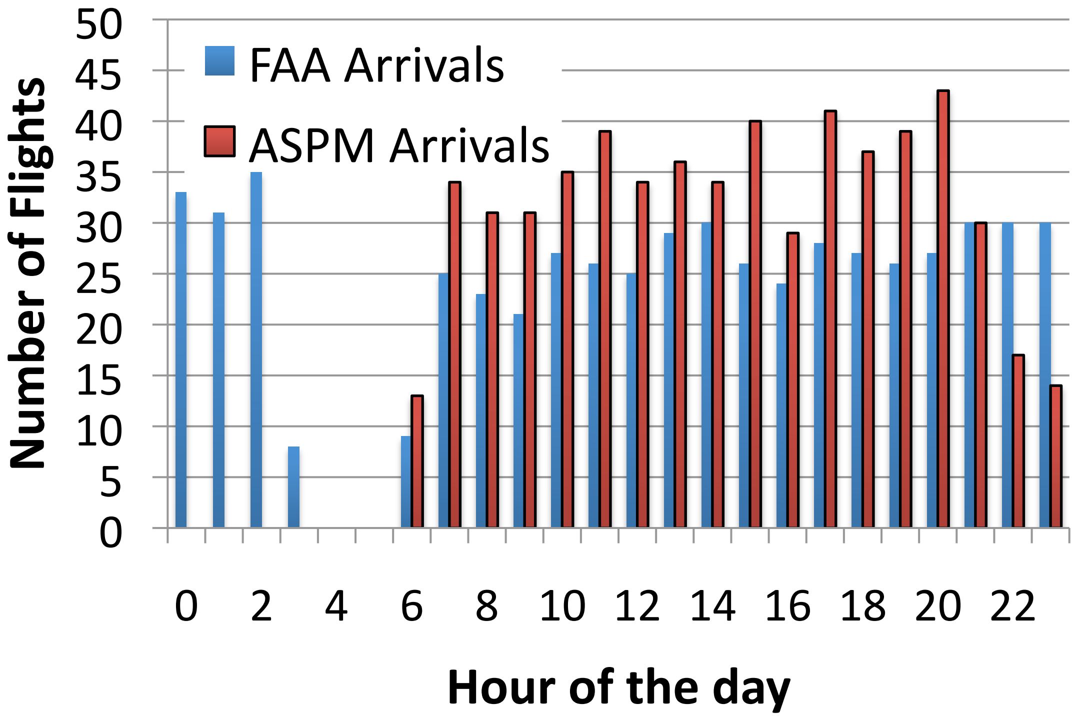



**Figure 52: # of flights per hour in 2010 at LGA**

### 4.7 CONCLUSIONS

The results of the medium fidelity analysis show that there are substantial potential benefits in the future from SCM at all of the 8 study airports. BOS has the lowest forecast benefits over the 20-year period, but still would realize over $30 million in fuel savings alone, or 3.8% of their taxi-out fuel costs. The analysis also showed that the future demand forecast could be overestimating future capacity and does not take gate constraints into account. Because the forecast was an input to the model that could not be replaced or modified, the resulting benefits estimate is highly uncertain. However, it is clear that at major airports such as JFK and ATL, even moderate increases in demand will result in severe congestion and that SCM shows the ability to reduce the environmental and fiscal impact of this congestion.

# 5. LOW FIDELITY METHODS FOR ASSESSING SCM BENEFITS AT NAS-WIDE AIRPORTS

## 5.1 BACKGROUND

As the results from the previous chapter have shown, potential benefits from departure metering can vary greatly from airport to airport. This makes extending the benefits assessment to a wider set of airports a challenge. We could not complete the medium fidelity analysis at all the OEP 35 airports in the time available for this thesis, but it would be a feasible and useful opportunity for future work. In order to make policy decisions, a fast, widely applicable model is desirable. Although these often imply low fidelity, they can be informed and validated by the medium and high fidelity models already presented. In addition, we recognize that an airport is a complex system and any individual low fidelity method is likely to deliver results that are substantially different from another model. Therefore, we present three different methods here to attempt to set bounds on the uncertainties in the estimation of benefits and to try to identify variables that have a strong correlation with benefits that can be used to identify airports where further study would be useful.

The first method weighs the benefits of the 8 medium fidelity study airports relative to the rest of the NAS using the relative amount of taxi delay as the indicator variable, i.e., the relative benefits expected at an airport scale with its relative taxi delay within the NAS.

The second method expands upon the first and replaces taxi delay with other variables that can be calculated for future years without simulation to create a linear regression. The regression was built using historical data from 2000 to 2010 and tested using benefits estimates for the OEP 35 airports in 2010.

The third method forms clusters of airports with similar characteristics that are assigned the mean benefit level of the medium-fidelity study airports in that set.

Each of these low fidelity modeling methods are examined in turn in this chapter. The recurring theme through this section will be the tradeoff between simplicity and ease of calculation on one hand and perceived accuracy of the results on the other. While all three methods fall under the umbrella of the ‘low fidelity’ method, there is substantial variation among them and each has their own merits and shortcomings.

## 5.2 METHOD 1: PERCENTAGE OF TAXI DELAY

Taxi delay should be an intuitive driver of SCM benefits, given SCM aims to address excess taxi times. Taxi delay at an airport can be calculated by finding the average unimpeded time (often taken as the 10$^{th}$ percentile of all taxi-out times) as well as the overall average time. The difference between the two multiplied by the total number of flights represents the taxi delay, or excess taxi time.

SCM is designed to reduce excess taxi time with engines on by spending that time held at a gate or other location. Therefore, one hypothesis is that taxi delay is linearly related to benefits from SCM. An assumption also needs to be made about how benefits and taxi delay relate into the future: the assumption used here is that the proportion of nationwide taxi delay at any given airport will remain constant into the future. Using this assumption, the percentage of delay at the 8 study airports from the medium fidelity methodology in 2010 is equal to the percentage of benefits both for the current day and into the future. From ASPM, the 8 study airports account for 40.8% of the taxi delay out of the OEP 35 airports in 2010, resulting in a scaling factor of 2.45. Therefore, one estimate of benefits from SCM at the OEP 35 airports is $8.9 billion if the unconstrained results are scaled, and $5.7 billion for the practical. Scaling the 2010 fuel estimate from [18] by the future demand, the total fuel burn in all stages of flight for departures from the OEP35 airports is 319 billion gallons, or $775 billion. The unconstrained benefits are then 1.1% of the total fuel cost, and the practical are 0.7%. Again, these numbers are probably overestimates.

This method has the advantage of being simple to calculate but suffers from the inaccuracy of a number of assumptions. The first is that taxi delay is directly proportional to benefits. Figure 53 and Figure 54 show that the benefits from SCM are much more concentrated at certain airports than the taxi delay. In particular, BOS, DFW and IAD have a substantially smaller proportion of SCM benefits than their proportion of taxi delay. This is related to the discussion on CDQM. There is a certain amount of taxi delay that is not affected by SCM, the 'target delay' of CDQM that allows airports to maintain maximum throughput. At BOS, DFW and IAD most flights that have taxi delays have delays that are under the target delay and therefore may not be as affected by SCM. Most airports with low taxi delay per flight will similarly have fewer benefits than their taxi delay suggests.

Another assumption that the method makes is that the ratio between the practical and unconstrained benefits remains the same across the OEP 35 airports as it was at the 8 study airports. The validity of this assumption is hard to gauge. Alternatively, the two estimates could be viewed as 'high' and 'low' estimates using the taxi delay method.

If the optimal target delay could be determined then it could be subtracted from the total taxi delay to find the possible benefits as in [5]. The problem, as mentioned before, is that the optimal target delay can vary from airport to airport. If the optimal target delay identified in the CDQM study of 6 minutes is applied to every airport, Figure 55 shows that the proportions still do not match well with the proportion of benefits; most notably, BOS, JFK and IAD have higher share than their share of benefits. Relative to the OEP 35, the 8 study airports now have 59.8% of the delay, which leads to a scaling factor of 1.67 and OEP 35 estimated benefits of $6.1 and $3.9 billion for the Unconstrained and Practical cases, which are 0.8% and 0.5% of total fuel cost.

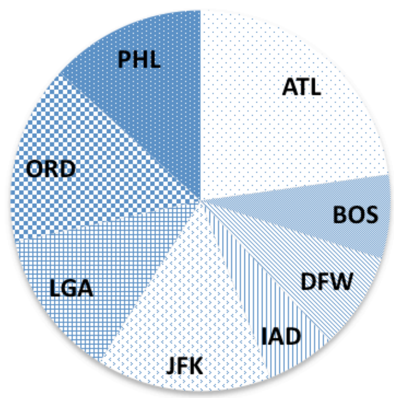



**Figure 53: 2010 Taxi Delay at 8 Study Airports**

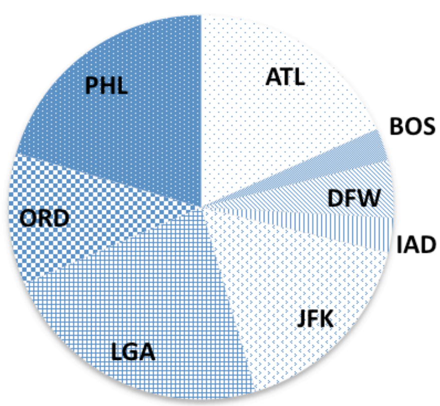



**Figure 54: 2010 SCM Benefits at 8 Study Airports**

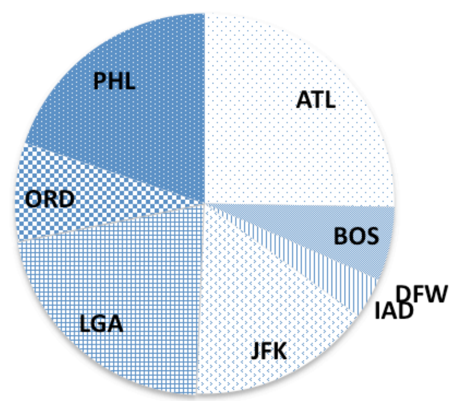



**Figure 55: 2010 Taxi Delay - 6 minutes at 8 Study Airports**

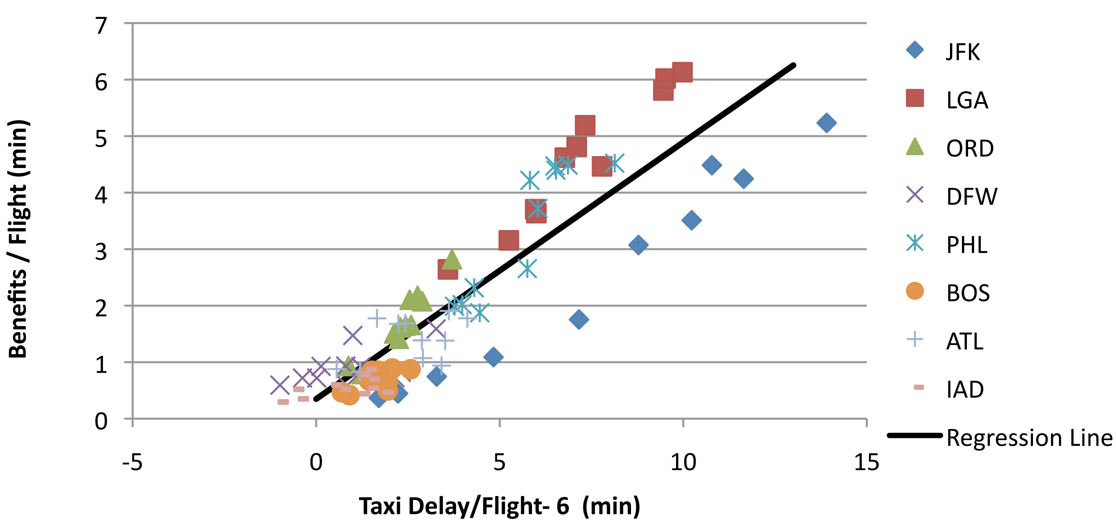



**Figure 56: Regression of Benefits to Reduced Taxi Delay**

The underlying assumption in this model is that taxi delay and SCM benefits are linearly related. A regression relating the amount of benefits from SCM per flight to the amount of modified taxi delay (delay – 6 minutes) per flight is shown in Figure 56. Using data from 2000 to 2010 results in the equation y = 0.4542x + 0.3491, with an $R^2$ of .7966. Most of the study airports are seen to lie on the regression line; however, Figure 56 shows that JFK is not well described by this equation. One hypothesis to explain this is that the target delay of 6 minutes is not as applicable at JFK as it is at other airports. Figure 57 shows that when the target delay for JFK is increased to 10 minutes, the $R^2$ is increased to .882 and the equation changes to y = .523x + .359. However, when this regression equation is used to predict the future benefits, there is substantial scatter between the Future (predicted by regression) and FutureActual (based on the results from the medium fidelity model), especially at higher delays. This highlights the problem that regressions should not be used to predict occurrences outside of the data they have been trained on. It also reflects the fact that this is not a fair comparison. The taxi delay calculated from the medium fidelity model does not include things like IMC conditions which have substantial impacts in actual operations.

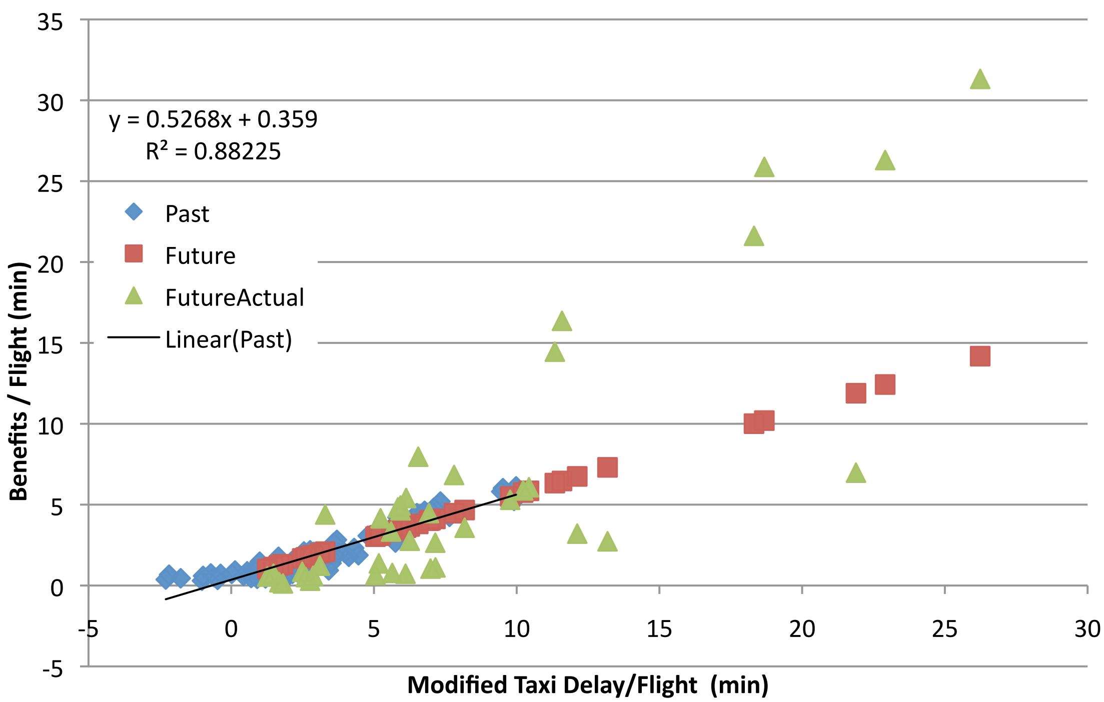



**Figure 57: Modified Taxi Delay Regression and Prediction of Future Benefits**

## 5.3 METHOD 2: MULTIPLE LINEAR REGRESSION MODEL

While the modified form of taxi delay shown in the previous section was shown to predict benefits from SCM quite well in the present, the performance was reduced in the future (as would be expected from a forecast). A more important problem is that taxi delay is only available for future years at the airports for which taxi models are available from higher fidelity models (i.e. which were studied in more detail). Because the goal of this chapter is to find a methodology that avoids extensive simulation, taxi delay is not a useful metric for a linear regression used for forecasting. New variables are needed that can be calculated for the future; the challenge is that this restricts the options to those variables that can be derived from the FAA forecast data; i.e., the pushback schedule and the forecasted capacity because taxi times are not available. 4 variables were chosen to represent congestion: total annual departure demand (# of pushbacks), total annual departure demand when an airport is operating at demand levels at or above capacity (# of pushbacks in congestion), the number of 15 minute periods in a year when the airport is operating at that point (periods with 100% usage), and the yearly average percentage of capacity used (% capacity used). For this analysis, a period of congestion was defined as a period where the number of pushbacks exceeded the departure capacity for that period. While this does not guarantee that the airport is actually operating at capacity (we do not know the condition of the runway), it was considered a reasonable proxy.

These four variables were chosen to represent different aspects of congestion and SCM. The total demand shows the overall size of the airport, because large airports would receive more benefits from an SCM system than a smaller airport with a similar delay per flight. The Pushbacks in Congestion variable represents how much of the total demand is operating when the airport is congested. This is different from the 100% Usage periods, which measures how often the airport is congested. Two airports that are congested for the same amount of time would have different benefits if the capacity is low at one airport (making the threshold for 100% usage low) often and the other airport is congested while in its highest capacity configuration. The % Capacity used variable gives an idea of the overall congestion at an airport.

To build the regression, data from the 8 medium-fidelity study airports from 2000 to 2010 was used. The benefits served as the dependent variable, and the four stated inputs were the independent variables. It was found that the percentage of capacity used does not significantly improve the quality of the model, so it was discarded. Several other statistical tests were performed to test the validity of the regression; these are listed and discussed in Appendix B. To mitigate the heteroscedasticity seen in the initial regression (the variance of the residuals was not constant, which violates one of the assumptions of a linear regression), the dependent variable (Benefits) was transformed by taking the square root. The resulting model is:

$$\sqrt{Benefits} = a_1X_1 + a_2X_2 + a_3X_3 + a_4$$

where $X_1$ is # of pushbacks, $X_2$ is periods of 100% usage, and $X_3$ is # of pushbacks in congestion.

The resulting parameters are shown below. As a check, data from the OEP 35 airports was obtained for the year 2010 and benefits were calculated using the saturation curve method as well as the linear regression. When the estimated benefits were compared to the actual benefits, the $R^2$ was actually higher (0.79) than the $R^2$ (.736, adjusted $R^2$ = .726) from the training data. Using data from the future schedules, the inputs were calculated and run in the model. Linear interpolations between the future schedule years were calculated to find the total benefits over the 2010-2030 period. The resulting benefits were $2.9 billion (0.3% of total fuel costs), much lower than the total calculated using the taxi delay method and only slightly more than the total practical benefits calculated for the 8 study airports. There is no distinction between the unconstrained and practical benefits in this method because the regression is built on historical data and the inputs in the future do not change.

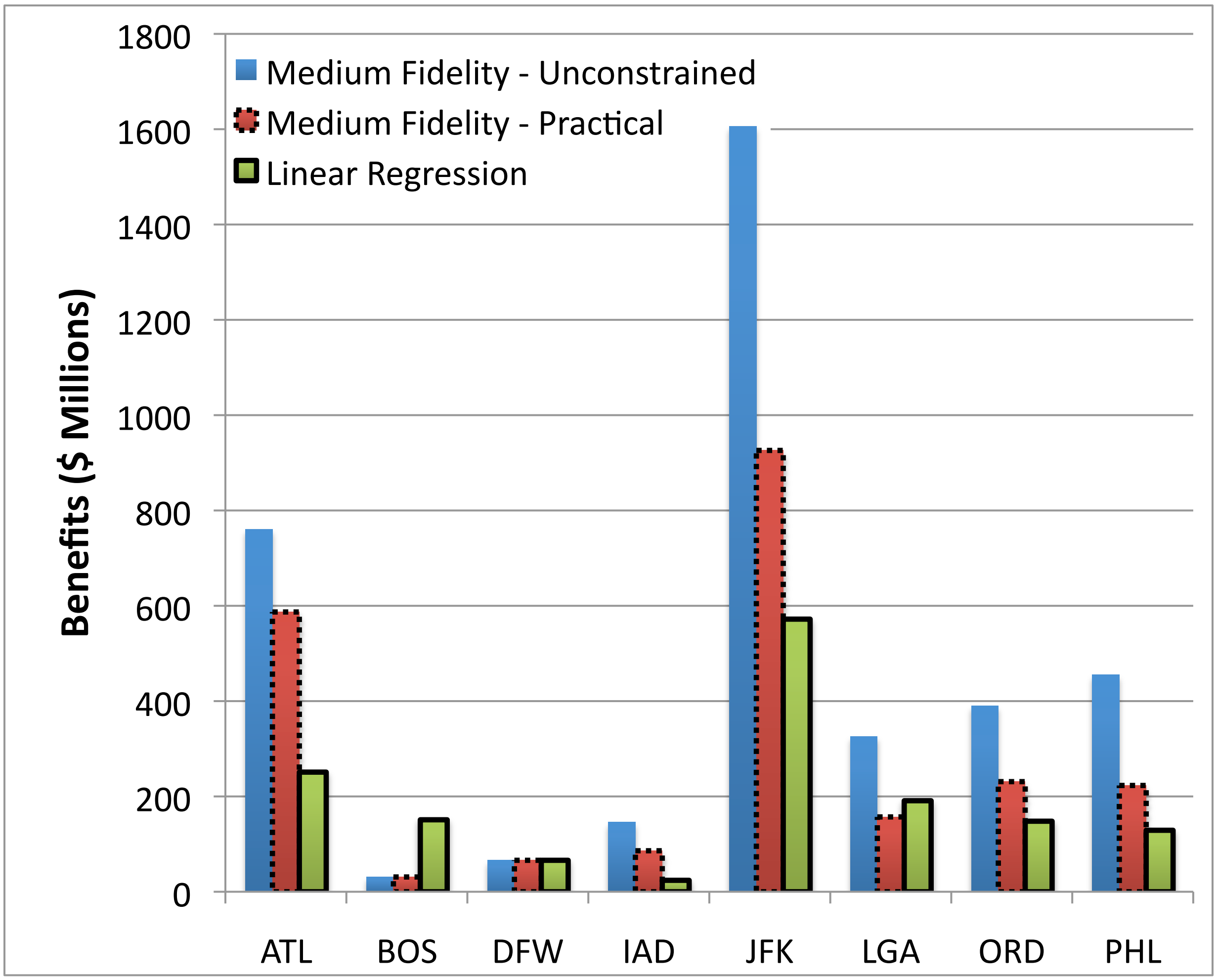



**Figure 58: Comparison of Medium Fidelity Method and Linear Regression**

Examining the individual estimates in Figure 58 (results for all airports given in Appendix B), the larger patterns are similar (JFK and ATL have the most benefits) but most airports have smaller benefits than in the medium fidelity study. There are some discrepancies; BOS and LGA both have more benefits than previously, and EWR, while not studied explicitly, is forecast to have benefits on the order of DFW (which is surprising, especially given the results in the clustering analysis discussed next). JFK and ATL were also underpredicted in the analysis of the residuals in Appendix B, and BOS was overpredicted. This continuation of previous trends is a check that the model is performing as expected. The lower total estimate reflects the fact that even with the square-root transformation, the regression still does not fully capture the dynamics of congestion.

**Table 5: Model Coefficients**

| Model | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|
| | B | Std. Error | Beta | | |
| Constant | -4.641 | 7.418 | | -0.626 | 0.533 |
| # of pushbacks | 19.292 | 3.142 | 0.662 | 6.14 | 0+ |
| periods of 100% usage | 14.257 | 1.467 | 1.004 | 9.717 | 0+ |
| # of pushbacks in congestion | -3.002 | 0.935 | -0.459 | -3.212 | 0.002 |

Despite the relatively high value of $R^2$, there is substantial cause for concern about the validity of the linear model. All three predictor variables can be considered positive indicators of congestion; i.e., the higher their value, the higher the congestion is expected to be. However, the coefficient for # of pushbacks in congestion is negative as shown in Table 5, indicating a decrease in congestion for an increase in the value of the variable. This suggests that instead of finding a model that truly describes congestion, we have only found the model that best fits the data given the low quality of the input variables. As a result, the future predictions should be considered even more uncertain than normal forecasts. If we calculate the residuals for the future predictions for the 8 study airports by comparing them to the Unconstrained forecast from the medium fidelity model and plot them with the residuals from the training data in Figure 60 (Original plot in Appendix B, rescaled in Figure 59), we see that the future residuals are both positively biased (large underpredictions) and much larger in magnitude than for the training data. This confirms our concerns that the linear model is not compatible with the medium fidelity model.

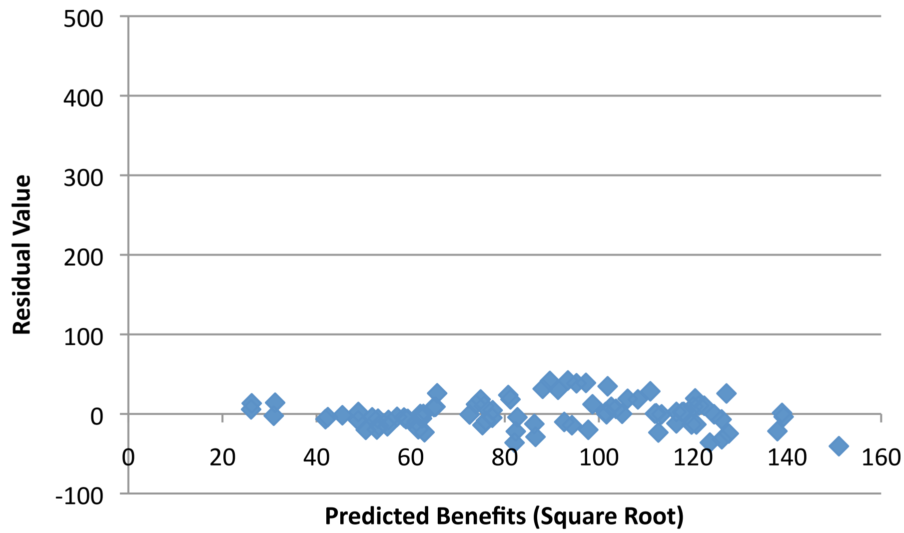



**Figure 59: Past Residuals Only**

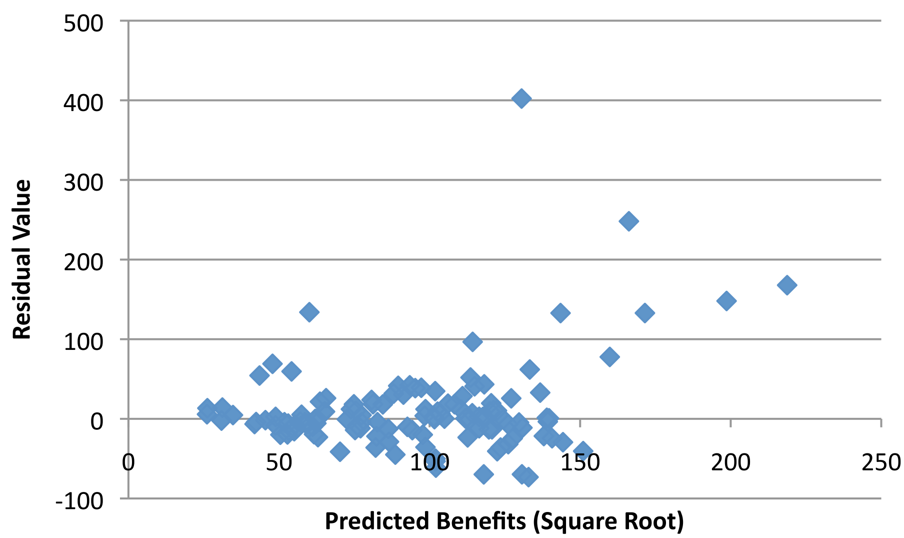



**Figure 60: Past and Future Residuals**

Given the concerns about the linear model as well as the observation that congestion increases nonlinearly as demand approaches capacity (Figure 3), a non-linear model might be expected to perform better. However it is not obvious what form a non-linear model should take.

Through experimentation this model was found to have an $R^2$ of .776, or slightly better than the linear model:

$$\ln(Benefit) = a_1\sqrt{Pushbacks} + a_2\exp(a_3 \times PeriodsOf100\%Use) + a_4 PushbacksInCongestion + a_5$$

When we use this equation to predict the future benefits, the estimate is increased to \$4.2 billion. However when we more closely examine the breakdown by airport in Figure 61, almost all of the increase is due to just one airport, JFK (with SFO increasing moderately). The benefits here are the cumulative benefits realized from 2010 to 2030 at each airport. This estimate for JFK is above even the unconstrained estimate from the medium fidelity method. It also highlights a concern about using a non-linear model: most of the historical data is in a range where benefits are still roughly linear. There is little data to show when and how an airport enters the non-linear regime. While the model correctly (according to the medium fidelity method) predicts JFK entering the non-linear regime, it misses ORD and ATL, shedding doubt on whether or not the estimate for SFO is accurate.

Overall, the nonlinear model adds much complexity and uncertainty while not providing corresponding additional insight or accuracy or even change in estimates from the linear model (apart from JFK). There is no guarantee that the form of the regression above is correct or optimal; the coefficient for Pushbacks in Congestion is still negative, as it was for the linear case. Finally, the nonlinear model makes the assumption that airports will tolerate operation in the nonlinear regime (akin to the Unconstrained case). In reality, they will probably try to avoid it as much as possible through regulations, schedule adjustments, or other forms of demand control. As a result, a non-linear method was not pursued any further.

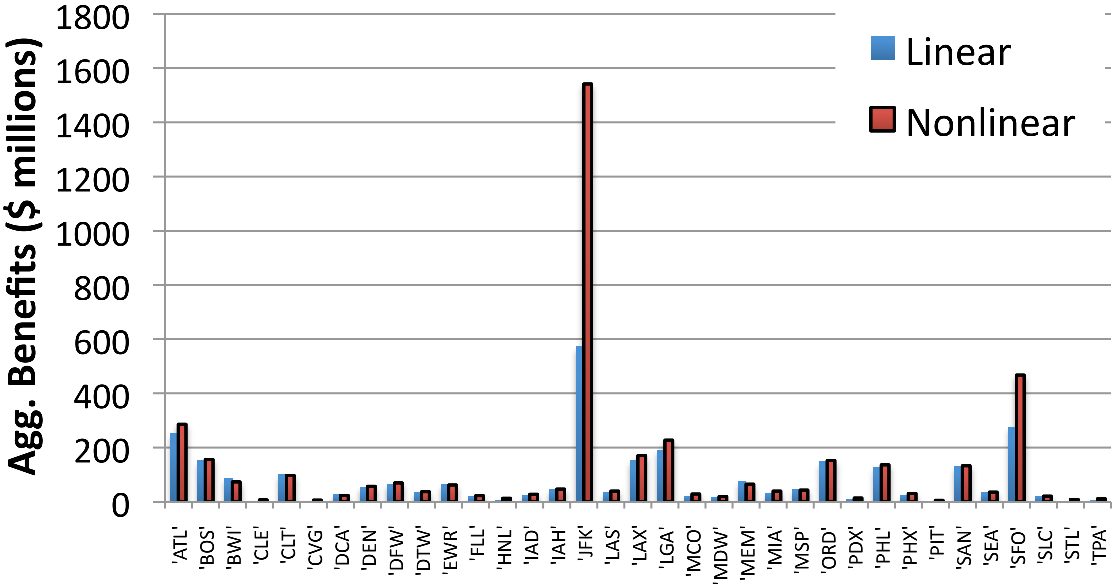



**Figure 61: Comparison of Linear and Nonlinear Methods**

## 5.4 METHOD 3: AIRPORT CLUSTERING

A third possible method is clustering, which groups airports into different bins based on certain key characteristics. The parameters chosen were variables that were identified as important based on the experience gained from the high and medium fidelity methods and were: Total Demand, Current % Capacity Used, Growth of Demand from 2010 to 2030 and Growth of Capacity from 2010 to 2030. The first two variables are the same in the linear regression and are present for the same reason as given previously. Clustering differs from the regression in that the evolution over time of the variables is not considered. As a result, a variable is needed to represent how traffic at an airport will change: the expected growth of congestion (Growth of demand and capacity). Clustering could be performed on these raw variables. However, given the insights from the high and medium fidelity methods, a new variable was created in order to create clusters that corresponded to historical levels of benefits. Several different equations were tested and the following both grouped the 8 study airports from the medium fidelity method into levels close to the practical benefits levels and also was relatively simple:

$$ClusterVariable = \left(\frac{DemandGrowth}{CapacityGrowth}\right) \times CapacityUsed^2 \times \sqrt{DemandLevel}$$

The form of the equation makes intuitive sense. It states that the most important variable is the capacity used (because it is squared) which agrees with Figure 3 that showed how congestion and delay increase nonlinearly as the percentage of capacity used nears 100%. Future congestion will also be affected by how much more the demand grows than the capacity. Finally, the overall level of traffic at an airport scales the benefits accordingly because the other two terms are percentages.

The clustering was performed using the k-means algorithm [24] with the clustering variable previously identified. The algorithm iteratively divides *n* observations into *k* clusters, where an observation belongs to the cluster whose mean it is closest to (Euclidean distance). *k* was chosen to be 4 for this analysis to both ensure multiple study airports be assigned to a cluster and to sufficiently stratify the results. 3 clusters (the Low, Medium and High levels) have at least one of the 8 study airports from the medium fidelity method, while Group 4 (Negligible) is estimated to have half of the benefits of BOS, the airport with the lowest level of benefit from the medium fidelity study set. Although Group 4 has no medium fidelity study airports, this was deemed to be acceptable because it consists of airports where metering is anticipated to be of little impact. Because all 8 study airports were chosen for their potential for SCM to have an impact, having one of them in the cluster would inflate the benefit estimate. To use clustering as a tool for generalizing benefits, the average benefits in terms of taxi time saved were found for the airports in a cluster. The remaining 27 airports from the OEP 35 were assigned to clusters based on their value of the clustering variable, and assigned the average value of the benefits at the study airports in that cluster. While airports are unique, we think this is a reasonable method because we are focusing on one specific attribute (congestion). It is also true that having small sample sizes of 2 or 3 airports providing the average benefit level can skew the results. This is partly alleviated because most of the airports are in the lower two tiers where the benefits are low, but future work should focus on airports such as EWR and SFO in the top tiers that have not yet been studied in detail. The resultant clusters and levels of benefit are shown below in Figure 62 and Table 6. JFK and ATL are the anchors for the “high” level of benefits, ORD, PHL and LGA are the medium level, and IAD, BOS and DFW are the low level, with no study airports in the negligible level. Most airports are in the low or negligible level clusters, as would be expected because most airports are not currently congested or operating close to their capacity. The airports not in the low or negligible tier are identified and basically conform to intuition.

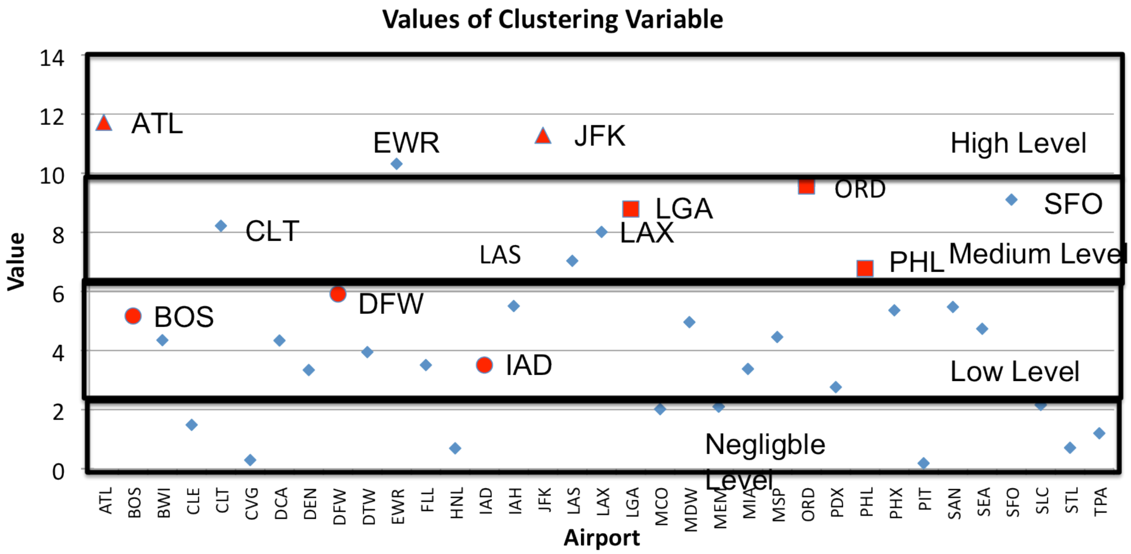



**Figure 62: Clustering Results**

**Table 6: Clustering Results**

| Unconstrained | | | |
|---|---|---|---|
| Cluster | Mean Benefit ($ Millions) | Number of Airports | Total Benefit ($ Millions) |
| High (JFK, ATL) | 1184 | 3 | 3551 |
| Medium (LGA, ORD, PHL) | 344 | 7 | 2407 |
| Low(BOS, DFW, IAD) | 81 | 16 | 1294 |
| Negligible (BOS/2) | 16 | 9 | 141 |
| **Totals** | | **35** | **7392** |

| Practical | | | |
|---|---|---|---|
| Cluster | Mean Benefit ($ Millions) | Number of Airports | Total Benefit ($ Millions) |
| High (JFK, ATL) | 757 | 3 | 2270 |
| Medium (LGA, ORD, PHL) | 204 | 7 | 1425 |
| Low(BOS, DFW, IAD) | 61 | 16 | 978 |
| Negligible (BOS/2) | 16 | 9 | 141 |
| **Totals** | | **35** | **4813** |

Table 6 shows the resulting benefits estimates from the clustering approach. Most of the 27 airports not studied in the medium fidelity method fall in the Low or Negligible category. The total unconstrained benefits (based on the average unconstrained benefits from the medium fidelity method) are estimated to be $7.4 billion (1.0% of total fuel cost) over the 20 year study period. The practical benefits are $4.8 billion (0.6 %). Just under half of the total benefits came from the three 'high-level' airports, ATL, EWR, and JFK.

## 5.5 CONCLUSIONS

Figure 63 shows that the three methods evaluated yielded results that varied substantially, from $3 to 9 billion, or between 0.4% and 1.2% of the total fuel cost across the OEP 35, as hypothesized (again, likely an overestimate of the percentage of total cost). The benefit estimate is the total benefits for the 2010-2030 period summed across the OEP 35 airports. We can attempt to explain the relative performance of the methods. The taxi delay scaling factor gave the largest estimate of benefits because not all taxi delay can in fact be saved by SCM. Airports with amounts of delay too small to be affected by SCM still receive benefits according to this model. By subtracting 6 minutes from the taxi delay, this effect is somewhat mitigated, but it is not as accurate as it could be because 6 minutes is not the optimal amount at every airport. The linear regression predicts much lower benefits than any of the scaling methods. The results are also the same for both the Practical and Unconstrained cases because physical constraints were not known for all 35 airports. As a result, no capping of the benefits was performed. Because many of the 8 study airports were predicted to approach or exceed their capacity in the medium fidelity model, their benefits increased non-linearly beyond any level seen in the past. The linear model could not capture this behavior and has correspondingly low benefits estimates. While a non-linear model could be constructed, it is not immediately obvious what its form should be. Also, given the low quality of the inputs it would probably not add any insights compared to the three current methods. Finally, clustering is, at its heart, a more intelligent scaling method so it makes sense that it is lower than the raw taxi delay but reasonably close to the modified taxi delay method.

The results from these methods also provide insight into which variables should be studied when choosing airports for further study or SCM implementation. While taxi delay is an intuitive choice, we have shown that raw taxi delay can overestimate the possible benefits. In addition, it is a hard metric to calculate in the future without detailed simulations. For long term planning, it is better to examine the relationship between the predicted growth in demand and the growth in capacity, as well as the current level of congestion (whether measured by taxi delay or

% of capacity used). An airport such as LGA which is currently congested but is forecast to remain at current traffic levels will receive fewer benefits when summed over the next 20 years than an airport which is not as congested but is forecast to face greatly increased demand, such as ATL.

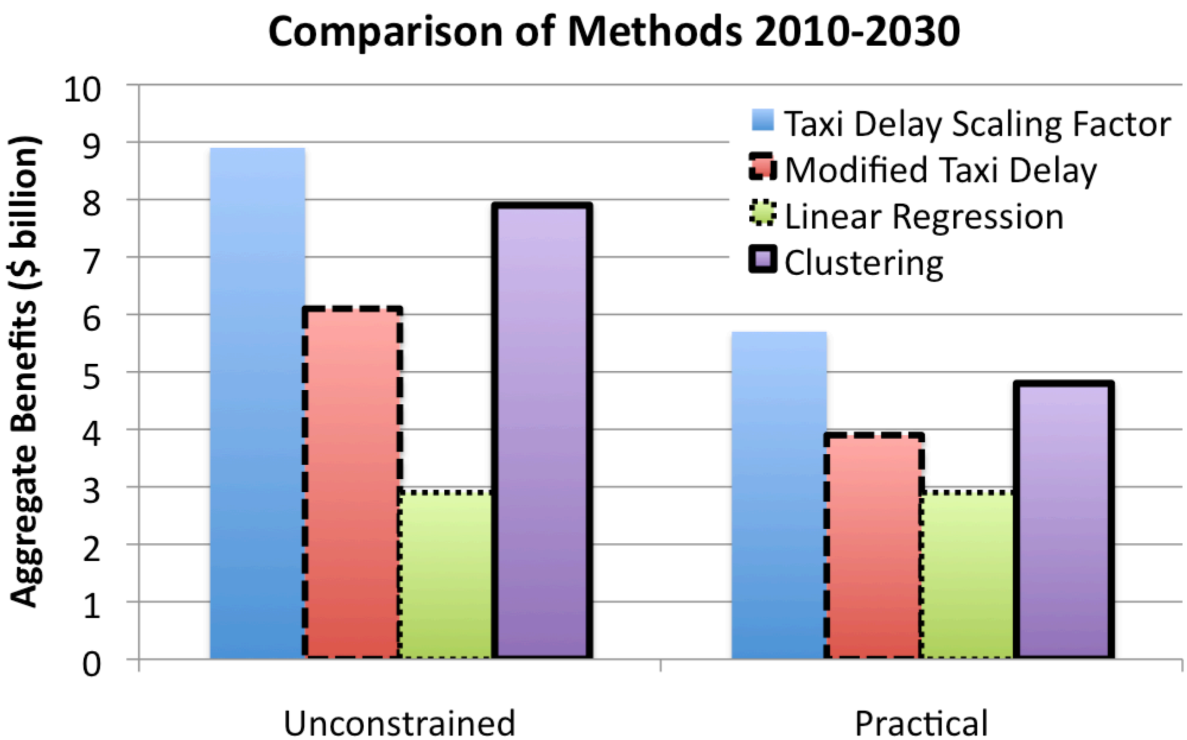



**Figure 63: Low Fidelity Methods**

# 6. CONCLUSIONS AND DISCUSSION

## 6.1 SUMMARY

Surface congestion management has already been shown to be an effective solution to the problem of congestion and excess taxi times at airports. However, in order to justify its expansion and inclusion in NextGen ATC systems, an analysis of the future benefits was required. This thesis showed that the benefits of such a system would be substantial, between \$3 and 9 billion over the next 20 years , or between 0.4% and 1.2% of the total fuel cost across the OEP 35 (again, likely an overestimate of the percentage of total cost). In addition, several airports currently have the potential for benefits on the level seen in the JFK field trial of \$10-20 million per year just in fuel costs, up to 10-20 % of the fuel costs for taxiing. By examining the predictors of congestion proposed in this thesis, policymakers can identify where to focus resources for maximum effect.

Three different methods were presented in this thesis to develop these future year SCM benefits estimates. The first, a high fidelity and highly specialized model of JFK airport in New York City, examined the change in a traffic metric (takeoff queue) and its related impact on taxi-out time to estimate the savings from implementing SCM. The model was based on ASDE-X surveillance data as well as ASPM taxi time information and looked at several secondary effects such as throughput, taxi in times, configuration-specific effects, and off-gate holding. The results showed that even with a substantial number of off-gate holds the overall efficiency of SCM (hold minutes to minutes of benefits) was still high, around 0.8 and resulted in expected benefits of between \$10 to 15 million annually.

This result informed the development of the medium fidelity approach to future benefits, where 8 airports were studied (including JFK). Configuration-specific effects were still examined as well as a lower fidelity method of examining the need for off gate holds, but other secondary effects were not included. The use of saturation curves to approximate whatever type of metering will be implemented in the future was validated by examining three field trials, each with a different type of SCM. The use of simulations to calculate airport performance and taxi times as well as the uncertainty in the future demand forecasts cause the fidelity of this method to be lower.

The final method was actually several different low-fidelity methods for calculating future benefits at many airports. While using any single method to perform a benefits analysis would not be recommended, examining them together helps to define the possible range of benefits as well as the uncertainty in an estimate of total benefits. In addition, particular airport characteristics were identified as relatively well-correlated with benefits. While the medium and high fidelity analyses have shown that airports are too complex to develop accurate overarching generalities, these characteristics are still useful for policymakers when trying to identify airports to implement SCM. Taken together, the suite of multi-fidelity models has been seen to be an effective method of undertaking a benefits assessment for applications such as SCM that can aid policy-making decisions.

## 6.2 IMPLICATIONS FOR POLICYMAKERS

Estimating the benefits of implementing a new technology or method serves two purposes. First, to decide whether or not it should be introduced and second, to determine where it should be introduced to provide the most benefit. While the first question cannot be answered without a corresponding analysis of the costs of implementation, this thesis has shown that the potential benefits of SCM are substantial (of the order of $ billions across the NAS, around 1% of total fuel costs and between 5 and 50% of taxi fuel costs, depending on the airport and year), especially considering the relative lack of infrastructure changes and improvements necessary. Depending on the deployment costs, this suggests SCM is likely to be an important contributor to future air transportation system enhancement. Chapter 5 addressed the second question by identifying variables and characteristics at airports that signal an opportunity for SCM to have a substantial impact.

The complexity of airport operations ensures that there is not one single variable that can accurately predict the usefulness of SCM at an airport. Taxi delay is an intuitive choice, but as was previously explained, there is a certain amount of taxi delay that SCM will not affect, given current operations. In addition, estimates of future delay at airports are highly uncertain. Instead, taxi delay and taxi times should be used as one of several indicators for present-day analyses, while future indicators could be the ratio of growth in demand to growth in capacity (with a higher value indicating more congestion), the daily demand level, both present and future (with a higher value indicating more absolute benefits because there are more flights that would be affected), and the percent of capacity that is used (with a higher value indicating more congestion). While JFK airport already has an SCM system in place, our analyses have shown that EWR, SFO, LGA, PHL and ATL would see substantial benefits from an implementation in the next 5 years, with the addition of ORD in the 10-20 year time frame. Depending on the cost of implementation, many other airports would also have measurable benefits.

Policymakers should also consider other changes that may be necessary both to implement SCM and to generally improve the performance of airports. The analysis of physical constraints at airports (availability of gates to perform gate-holding at) assumed that gates were interchangeable when in fact airlines “own” specific gates. If this policy could be changed so that gates are a common resource it would better support SCM strategies as well as other potential improvements such as accommodating new aircraft with increased wingspan.

A related issue is that of gate space. As was shown in the medium fidelity method, several of the study airports are forecast to face demand for gates that greatly exceeds the current capacity. While some airports such as ORD have plans for such a contingency, other airports such as JFK that are space constrained may not be able to accommodate all the holds SCM may require (and possibly may not be able to accommodate the forecast demand). The resulting 'cap' on the demand has large impacts on both the benefits from SCM (and virtually every other proposed improvement) and planning for the US air transportation system in general. Policymakers should examine the implications of lower demand levels than currently forecast as well as work to find ways to possibly accommodate increased demand at constrained airports.

SCM is just one of many improvements being considered for future implementation. Many of the assumptions, considerations and limitations studied in this thesis are also relevant in assessing other improvements. The most important is probably ensuring that the demand forecast is as accurate as possible. As was shown in the medium and low fidelity analyses, the demand forecast has a substantial impact on the benefits of SCM. While congestion is strongly tied to the demand, almost every improvement such as choosing the correct runway configuration or departure sequencing also depends on the demand. Accounting for some of the micro-level constraints such as gate availability or redistribution of flights to off-peak times that are not always considered when making system-wide forecasts is vital for when studying the benefits.

The other broadly applicable lesson from this thesis is that airports are unique. While we presented several methods for generalizing the benefits, they have substantial uncertainty and should be treated only as order-of-magnitude approaches. In addition, we examined several cases of large changes in airport behavoior or performance, such as the variation of performance at DFW where only a detailed investigation into airport specifics yielded the probable cause. Obviously such a detailed analysis is not possible at every airport or for every study, but generalization techniques should be applied with great care. The clustering performed in Chapter 5 is a good example. The clustering variable was picked to reflect SCM, and the resultant groups are meaningful only in the SCM context. Trying to use them for studying departure sequencing, for example, is not recommended.

## 6.3 UNCERTAINTY

As might be expected in systems as complex and dynamic as airports, the uncertainty in our predictions of SCM benefits is high. Even in studies using historical operational data, such as the high fidelity analysis of JFK, there are many sources of uncertainty. Some were explicitly assessed, such as the range of possible gate-to-spot correction factors (5-9 minutes). Others, such as the precise location of off-gate holds and the interaction between SCM, maintenance, ground holds, EDCT's, and other causes of ground delay were not addressed. In the future analyses, the inputs alone have large uncertainties. The demand for air travel is forecast to steadily rise over the next 20 years, but history has shown that constant growth is rarely the case. Recessions, wars, terrorist events, fuel prices and more all can have negative impacts on the demand. Economic booms, revolutionary technology, and low fuel prices could all increase the demand beyond the forecast. Airports could decide that a planned expansion could cost too much or that a new runway is suddenly necessary, throwing off capacity estimates. Changes in regulations, procedures, or fleet mix could lead to substantial changes in airport performance as was hypothesized to have happened at DFW.

The source of uncertainty with the largest effect on benefits is the future demand. The current forecasts do not consider physical constraints and have optimistic capacity forecasts. As a result, the future demand will most likely be lower than these forecasts. All of these issues need to be explicitly identified so their implications can be explicitly considered when the results are being interpreted. A very pessimistic lower bound on benefits can be calculated by assuming that demand remains at 2010 levels. In this case, the 8 study airports from the medium fidelity method would have the \$38.4 million in benefits from 2010 in each subsequent year to give \$806 million over the period from 2010 to 2030, 0.1% of total fuel costs. This is substantially smaller than the \$2.79 billion in benefits in practical benefits previously estimated.

## 6.4 FUTURE WORK

There are several areas where future work would be useful. Developing a forecast of demand that explicitly accounts for more of the key actual or expected constraints within the system would greatly decrease the uncertainty in the analysis, but would require a more detailed set of models. In addition, the traffic patterns may change as an airport nears both its runway and terminal capacity. It was shown in Chapter 4 that the model currently scales traffic each year. Instead, traffic may move to lower-demand periods to try to escape congestion. Changing the forecast model, however, is likely an enormous task by itself. Another possibility would be to examine multiple forecasts (low, medium, high) to better develop a range of estimates. A human-in-the-loop simulation of future traffic scenarios would help to validate the estimates of congestion as well as provide an estimate of the benefits possible from SCM. This would be especially valuable given that many airports will face levels of traffic that have never been seen before. Observing how controllers react to these traffic levels would be very helpful in predicting future performance. More work could also be done on comparing the different methods of SCM; which methods work best at which airports, and how the levels of benefits compare across methods in similar situations. Finally, further study on airports where the performance has changed substantially (DFW) or is forecast to change (ORD) would be informative. Interviews with controllers, planners, and other airport staff would be helpful in building a more complete model of airport performance both in the past and the future.

# 7. REFERENCES


[1] Federal Aviation Administration (FAA), Aviation System Performance Metrics (ASPM) database, http://aspm.faa.gov/, 2011.

[2] de Neufville, R., A. Odoni, Airport Systems Planning Design and Management, New York, NY, 2003

[3] FAA Office of Policy and Plans (APO), Terminal Area Forecast (TAF) data, retrieved from http://aspm.faa.gov/main/taf.asp, 2009.

[4] Simaiakis, I., H. Balakrishnan, H. Khadilkar, T. G. Reynolds, R. J. Hansman, B. Reilly, & S. Urlass, "Demonstration of Reduced Airport Congestion Through Pushback Rate Control", *9th USA/Europe ATM Research & Development Seminar* (ATM2011), Berlin, Germany, Jun. 2011.

[5] Brinton, C., C. Provan, S. Lent, T. Prevost & S. Passmore, "Collaborative Departure Queue Management: An Example of Collaborative Decision Making in the United States", *9th USA/Europe ATM Research & Development Seminar* (ATM2011), Berlin, Germany, Jun. 2011.

[6] Nakahara, A., T. G. Reynolds, T. White, C. Maccarone & R. Dunsky, "Analysis of a Surface Congestion Management Technique at New York JFK Airport", *11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, 2011.

[7] Eurocontrol, Airport CDM Operational Concept Document, Version 3.0, September 2006

[8] Jung, Y., T. Hoang, J. Montoya, G. Gupta, W. Malik & L. Tobias, "Performance Evaluation of a Surface Traffic Management Tool for Dallas/Fort Worth International Airport", *9th USA/Europe ATM Research & Development Seminar* (ATM2011), Berlin, Germany, Jun. 2011.

[9] Simaiakis, I., "Modeling and Control of Airport Departure Processes for Emissions Reduction", Master's Thesis submitted to Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, http://hdl.handle.net/1721.1/58289, 2009.

[10] Nicolas Pujet. "Modeling and control of the departure process of conjested airports". PhD thesis, Massachusetts Institute of Technology, 1999.

[11] Robert Arthur Shumsky. “Dynamic statistical models for the prediction of aircraft take-off times” PhD thesis, Massachusetts Institute of Technology, 1995.

[12] Bhadra, D., D. Knorr & B. Levy, “Benefits of Virtual Queuing at Congested Airports Using ASDE-x: A Case Study of JFK Airport”, *9th USA/Europe ATM Research & Development Seminar* (ATM2011), Berlin, Germany, Jun. 2011.

[13] Stroiney, S., B. Levy, C.J. Knickerbocker, “Departure Management: Savings in Taxi Time, Fuel Burn, and Emissions”, AIAA Integrated Communications, Navigation and Surveillance (ICNS) Conference, Washington DC, 2010.

[14] Stroiney, S. & B. Levy, “Departure Queue Management Benefits Across Many Airports”, AIAA Integrated Communications, Navigation and Surveillance (ICNS) Conference, Washington DC, 2011.

[15] Ball, M. O., C.-Y. Chen, R. Hoffman & T. Vossen, “Collaborative Decision Making in Air Traffic Management: Current and Future Research Directions”, in New Concepts and Methods in Air Traffic Management, L. Bianco, P. Dell’Olmo & A. R. Odoni (Eds.), Springer, 2001.

[16] H. Idris, J. P. Clarke, R. Bhuva, and L. Kang. Queuing model for taxi-out time estimation. *ATC Quarterly,* 2001.

[17] International Civil Aviation Organization, “ICAO Engine Emissions Databank,” July 2010, http://www.caa.co.uk/default.aspx?catid=702, 2011.

[18] Bureau of Transportation Statistics, Airline Fuel Cost and Consumption (U.S. Carriers - Scheduled) January 2000 - May 2011, http://www.transtats.bts.gov/fuel.asp [accessed 20 May 2012].

[19] I. Simaiakis and N. Pyrgiotis, “An Analytical Queuing Model of Airport Departure Processes for Taxi Out Time Prediction”, *10th Annual AIAA ATIO Conference*, 2010.

[20] Leo Breiman, “Random Forests”, *Machine Learning*, 45(1):5–32, 2001.

[21] Kimberly Noonan, “FAA’s System-Wide Analysis Capability”, *Integrated Communications, Navigation and Surveillance Conference (ICNS), 2011* , vol., no., pp.1-28, 10-12 May 2011 doi:10.1109/ICNSURV.2011.5935365 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5935365&isnumber=5935117

[22] FAA/MITRE Corporation, "Capacity Needs in the National Airspace System 2007-2025: An Analysis of Airports and Metropolitan Area Demand and Operational Capacity in the Future", FACT2 Report, 2007.

[23] FAA, “Economic Information for Investment Analysis Data Package”, March 2011.

[24] MacQueen, J. B. (1967). "Some Methods for classification and Analysis of Multivariate Observations", Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. pp. 281–297.

[25] FAA, Airport Diagrams. URL: http://www.faa.gov/airports/runway_safety/diagrams/

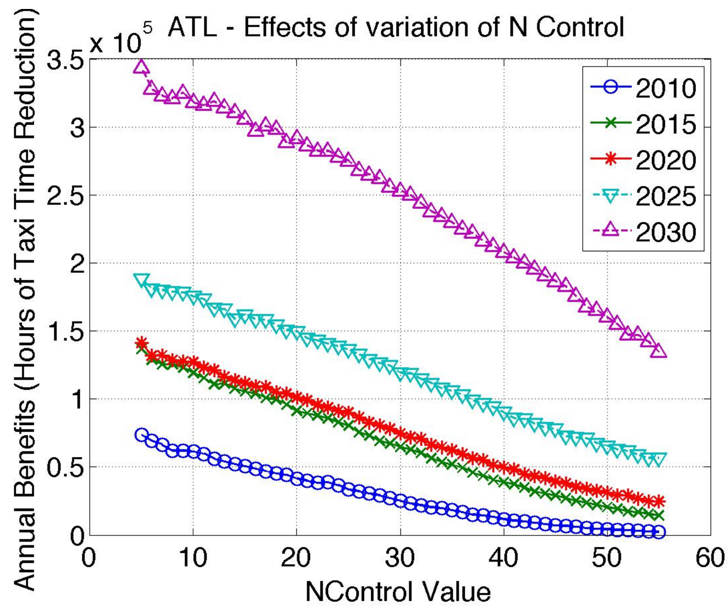
## APPENDIX A: ADDITIONAL INFORMATION

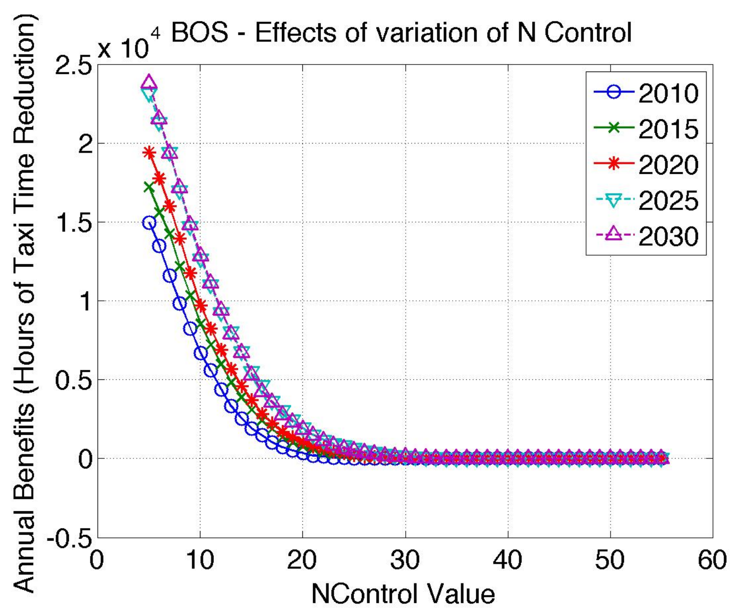
### Random Forest Variables

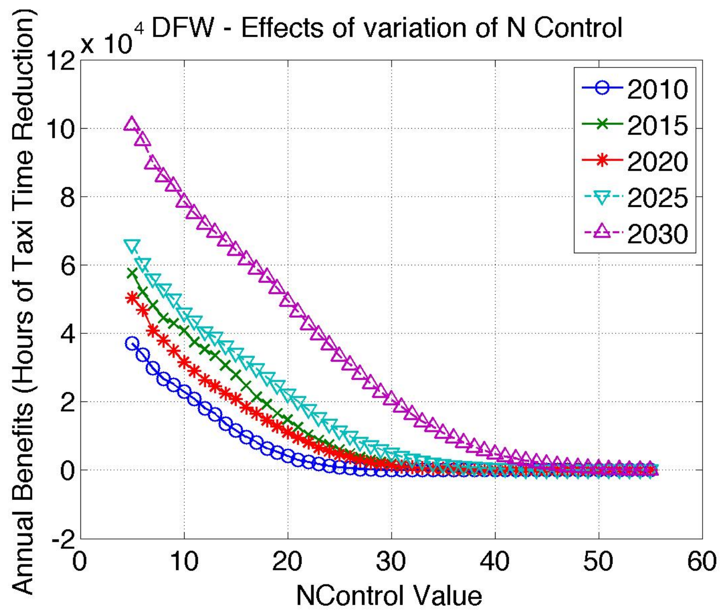
| Variable Name | Description | Source |
|---|---|---|
| Mean(DepDemand) | Mean(depDemand) - Mean Departure Demand (Yearly, by configuration, per hour) | ASPM - APM |
| Mean(arrDemand) | Mean(arrDemand) - Mean Arrival Demand (Yearly, by configuration, per hour) | ASPM - APM |
| 90%DepDemand | 90%DepDemand - 90th percentile Departure Demand (Yearly, by configuration, per hour) | ASPM - APM |
| 90%ArrDemand | 90%ArrDemand - 90th percentile Arrival Demand (Yearly, by configuration, per hour) | ASPM - APM |
| Mean(depCap) | Mean(depCap) - Mean Departure Capacity (Yearly, by configuration, per hour) | ASPM - APM |
| mean(arrCap) | mean(arrCap) - Mean Arrival Capacity( Yearly, by configuration, per hour) | ASPM - APM |
| 90%DepCap | 90%DepCap - 90th percentile Departure Capacity (Yearly, by configuration, per hour) | ASPM - APM |
| 90%ArrCap | 90%ArrCap - 90th percentile Arrival Capacity (Yearly, by configuration, per hour) | ASPM - APM |
| Used | Used - % of Configuration | ASPM - APM |

| | | |
|---|---|---|
| | Capacity Used (Yearly, by configuration) | |
| # of Arrival Runways | # of Arrival Runways (By Configuration) | Derived from Configuration |
| # of Departure Runways | # of Departure Runways (By configuration) | Derived from Configuration |
| # of Unique Departure Runways | # of Unique Departure Runways (By configuration) | Derived from Configuration |
| # of Unique Runways | # of Unique Runways (By Configuration) | Derived from Configuration |
| Area of Airport | Area of Airport (acres) | Wikipedia / Airport Websites |
| Miles of Taxiway | Miles of Taxiway (total) | Wikipedia / Airport Websites |
| Miles of Runway | miles of Runway (total) | Wikipedia / Airport Websites |
| Terminals | Terminals | Wikipedia / Airport Websites |
| Traffic | Traffic - Total traffic at the airport (Yearly) | ASPM-APM |
| %Capacity used | % Capacity Used - % of Airport capacity used (Yearly) | ASPM-APM |
| Gates | # of Gates | Wikipedia / Airport Websites |
| Nstar | Saturation point of Configuration | From Simaiakis code - remove top 2.5% of flights (by N), find the N value where the throughput reaches 95% of maximum throughput |
| ThS | Saturation Throughput - throughput at saturation point | From Simaiakis code |

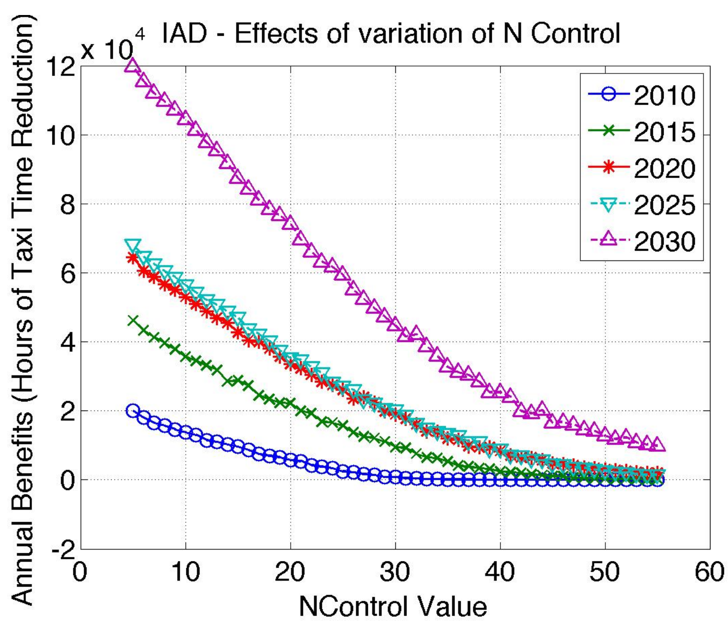
## Variation of benefits with N control



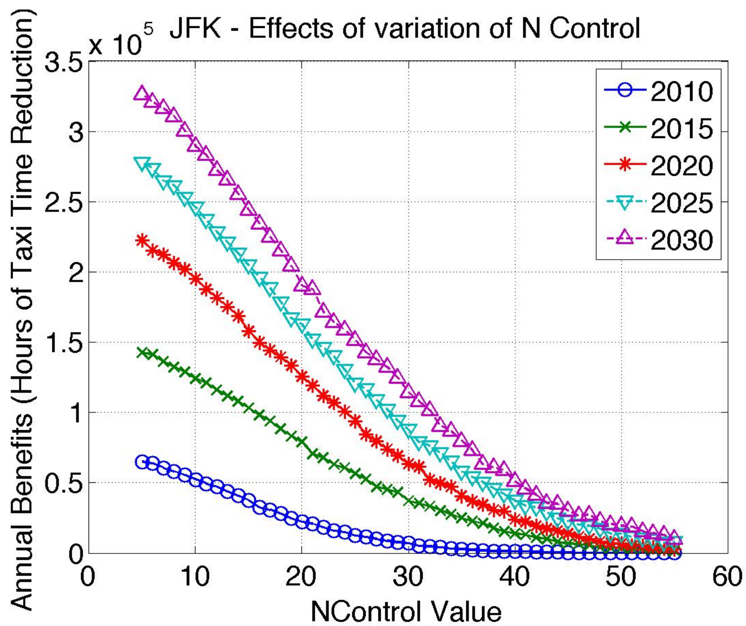



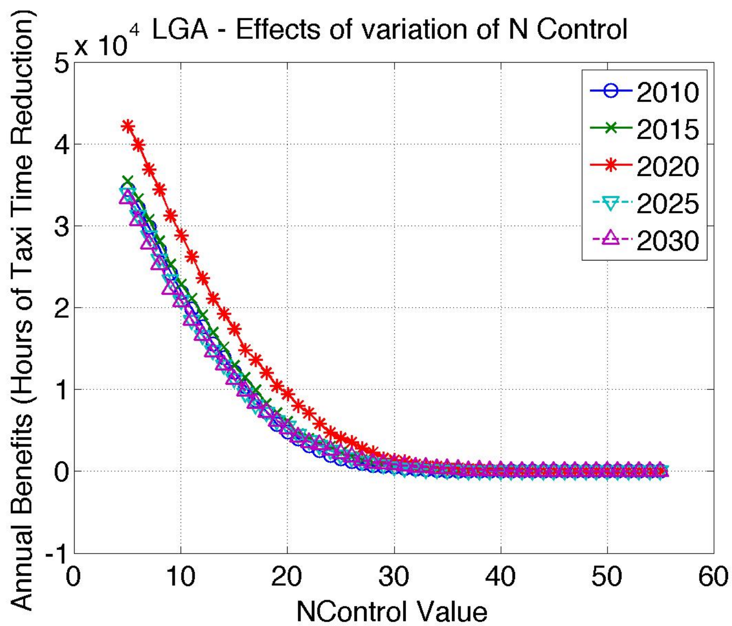


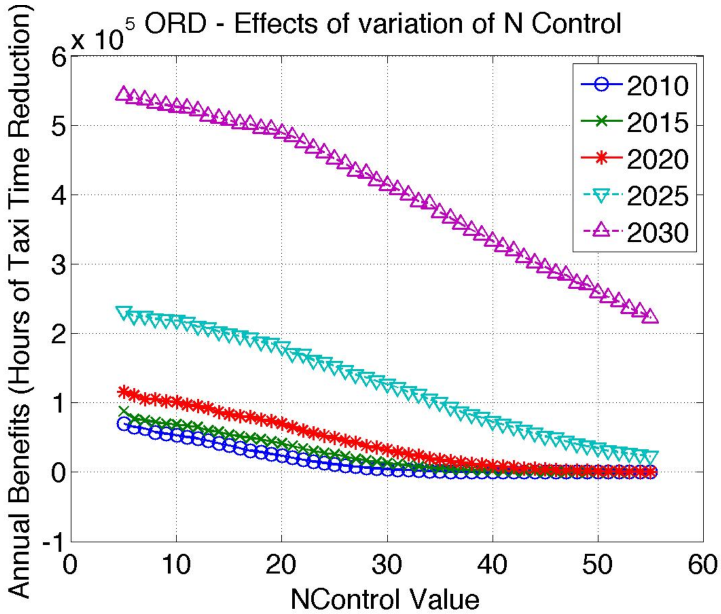
DFW - Effects of variation of N Control
x 10^4
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030
12
10
8
6
4
2
0
-2
0
10
20
30
40
50
60

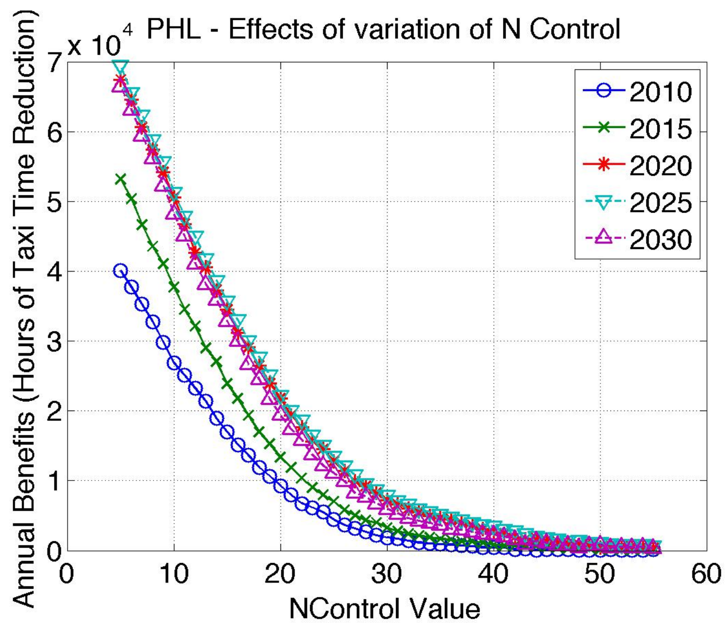



IAD - Effects of variation of N Control
x 10^4
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030
12
10
8
6
4
2
0
-2
0
10
20
30
40
50
60


JFK - Effects of variation of N Control
x 10^5
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030


LGA - Effects of variation of N Control
x 10^4
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030


x 10^5 ORD - Effects of variation of N Control
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030
6
5
4
3
2
1
0
-1
0
10
20
30
40
50
60


x 10^4 PHL - Effects of variation of N Control
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030
7
6
5
4
3
2
1
0
0
10
20
30
40
50
60

## APPENDIX B: STATISTICAL TESTS

Several statistical tests were conducted to show the validity of using a linear regression model. The assumptions that were checked were for constancy of error variance (homoscedasticity), normality, collinearity, and the appropriateness of the model (whether a better form would fit the data, whether variables could be being omitted).

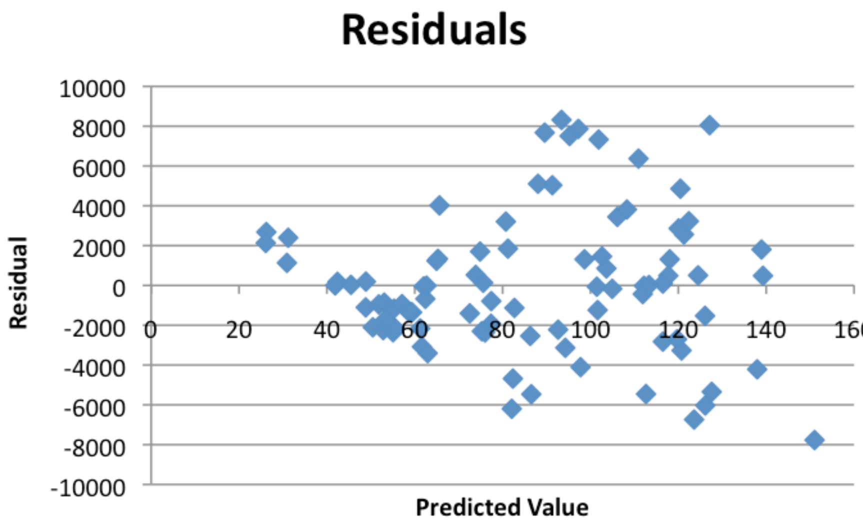



**Figure 64: Residual Plot, Initial Model**

Figure 64 shows the residuals from the initial model (using raw benefits as the dependent variable). When the residuals were plotted, the magnitude increased with the predicted value, showing heteroscedasticity. To compensate for this, the dependent variable was transformed by taking the square root.

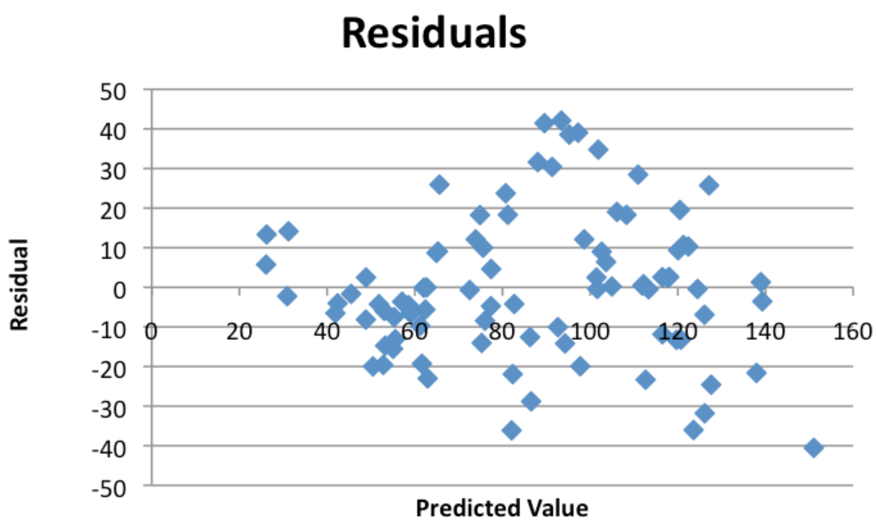



**Figure 65: Residual plot, transformed model**

While this only moderately improved the situation, the adjusted $R^2$ value did increase from .666 to .726, indicating a better fit to the data. The Breusch-Pagan test was run to test for constancy of the error variance where the test statistic is $\chi^2_{BP} = \frac{SSR*}{2} \div \left(\frac{SSE}{n}\right)^2$. SSR* is the regression sum of squares when regressing the squared residual on an individual predictor variable. We run three tests, one for each variable and compare to the critical value. At an a level of .05, the critical value is $\chi^2(0.95;1) = 3.84$.

**Table 7: Breusch-Pagan test values**

| | SSR* | $\chi^2_{BP}$ |
|---|---|---|
| Pushbacks | 51035 | 0.248 |
| Pushbacks in Congestion | 470882 | 2.287 |
| Periods with 100% Usage | 1873903 | 9.100 |
| | | |
| SSE | 28237 | |

$H_0$: $\gamma_1 = 0$ (There is no dependence on $X_i$ of the squared residual)

$H_a$: $\gamma_1$ does not equal 0 (The squared residual does depend on $X_i$)

If $\chi^2_{BP} > \chi^2(0.95;1)$ conclude $H_a$. Else conclude $H_0$

We can conclude $H_0$ only for Pushbacks and Pushbacks in Congestion (P values of 0.62 and 0.13), but we conclude $H_a$ for 100% Usage (P = 0.003). While this is not optimal, the model is not intended to provide accurate forecasts, only a general estimate and this is therefore acceptable.

In addition, a normal probability plot was made to assess whether the error terms were normally distributed. The coefficient of correlation was found to be r = .994. The critical value for n = 88 is 0.986, so there is support for our conclusion that the errors are distributed normally.

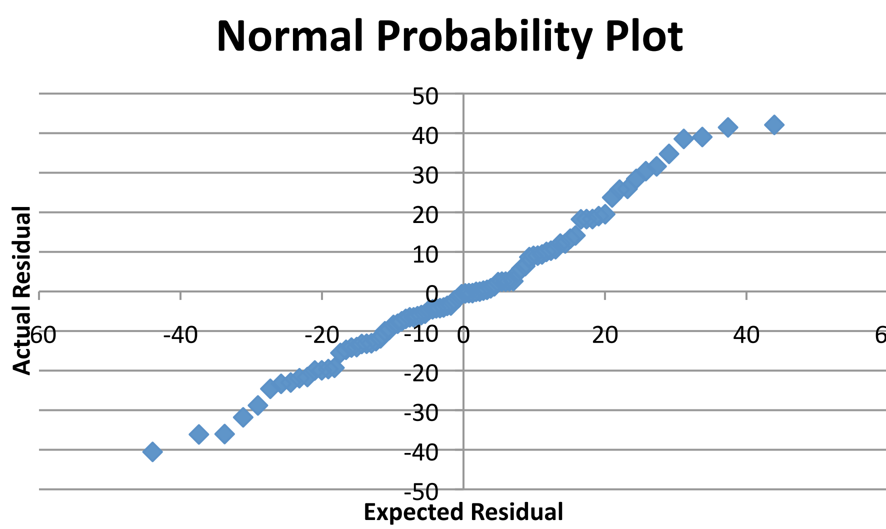



**Figure 66: Normal Probability Plot**

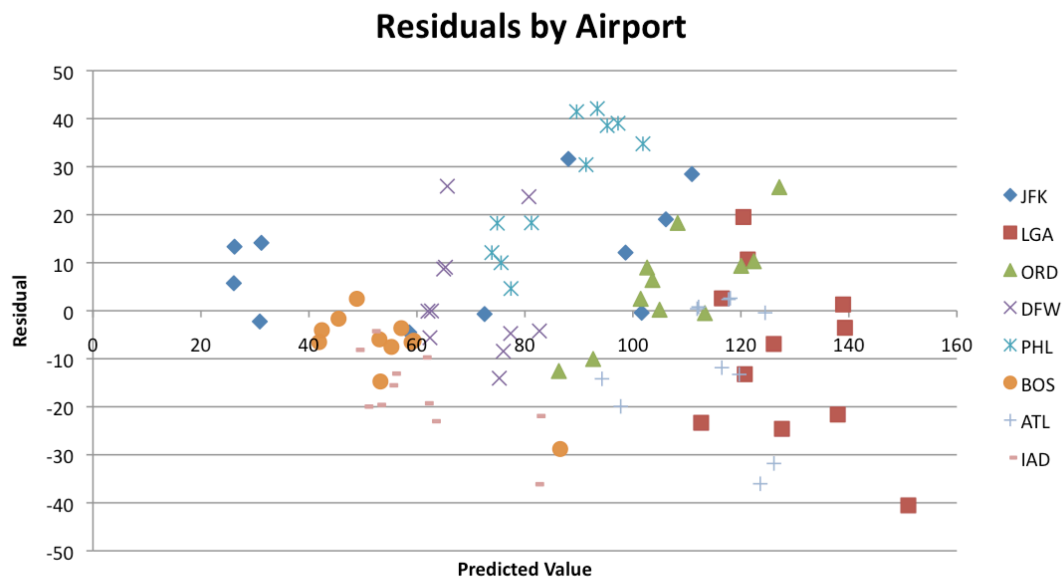



**Figure 67: Residuals categorized by airport**

When the residuals are broken down by airport, it is obvious that there are strong airport-specific effects, important variables that have been left out, or both. However, because extensive data is available only for these 8 airports, airport-specific effects cannot be quantified. While in general it seems like larger, more congested airports (JFK, ORD, PHL) are underpredicted (positive residuals), there are exceptions like LGA. There are also important variables that could be included to improve the results, such as taxi delay or average taxi time, but the model is limited to variables available in the future. Finally, there are definitely issues with collinearity among the independent variable, as shown in Figure 68, because they are all based on the same metrics, implying that using the model for predictive purposes may perform poorly. A formal method of detecting m get liulticollinearity is the Variance Inflation Factor (VIF). This can be calculated for each variable and describes how an independent variable is related to the other independent variables by the equation $VIF_k = \frac{1}{1 - R_k^2}$

**Table 8: Variance Inflation Factors**

| | VIF |
|---|---|
| Pushbacks | 3.61 |
| Pushbacks in Congestion | 6.33 |
| Periods with 100% Usage | 3.32 |

Typically values over 10 indicate severe multicollinearity. While the VIF here are relatively high, they do not exceed 10 and multicollinearity is not considered further. There are interesting relationships that appear to be separated based on the number of runways in use (1 or more than 1) but because this is not an experiment where the conditions can be controlled and that almost all variables at an airport are inter-related, not much can be done about this. Given these restrictions, the large uncertainty present in the inputs, and the results from the Breusch-Pagan and normality tests, the model was deemed to be sufficient for the purposes of this thesis.

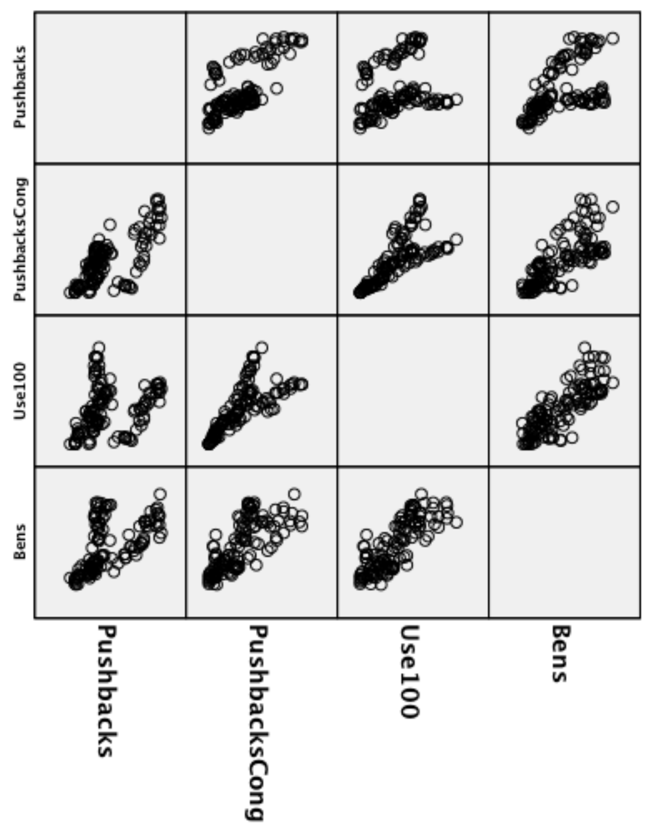



**Figure 68: Scatter plots of Variables**

**Table 9: Multiple Regression Results**

| | Thousand Hours Taxi Time Reduction | Million Gallons | $ Millions |
|---|---|---|---|
| ATL | 413 | 103 | 251 |
| BOS | 284 | 62 | 151 |
| BWI | 142 | 36 | 88 |
| CLE | 7 | 1 | 3 |
| CLT | 227 | 41 | 101 |
| CVG | 5 | 1 | 2 |
| DCA | 60 | 12 | 28 |
| DEN | 108 | 23 | 55 |
| DFW | 106 | 27 | 66 |
| DTW | 70 | 15 | 37 |
| EWR | 107 | 27 | 64 |
| FLL | 34 | 8 | 20 |
| HNL | 6 | 2 | 4 |
| IAD | 50 | 10 | 24 |
| IAH | 102 | 19 | 47 |
| JFK | 655 | 235 | 572 |
| LAS | 55 | 14 | 35 |
| LAX | 203 | 63 | 153 |
| LGA | 397 | 79 | 191 |
| MCO | 32 | 9 | 22 |
| MDW | 30 | 7 | 17 |
| MEM | 106 | 32 | 77 |
| MIA | 39 | 13 | 33 |
| MSP | 91 | 19 | 46 |
| ORD | 250 | 61 | 148 |
| PDX | 19 | 4 | 10 |
| PHL | 264 | 53 | 129 |
| PHX | 45 | 11 | 26 |
| PIT | 2 | 0 | 1 |

| | | | |
|---|---|---|---|
| **SAN** | **221** | **55** | **133** |
| **SEA** | **56** | **14** | **35** |
| **SFO** | **371** | **114** | **276** |
| **SLC** | **44** | **8** | **20** |
| **STL** | **6** | **1** | **3** |
| **TPA** | **9** | **2** | **5** |
| **Total** | **4615** | **1182** | **2872** |

**Table 10: Airport Clusters**

| Airport | Group |
|---|---|
| ATL | 1 |
| EWR | 1 |
| JFK | 1 |
| | |
| CLT | 2 |
| LAS | 2 |
| LAX | 2 |
| LGA | 2 |
| ORD | 2 |
| PHL | 2 |
| SFO | 2 |
| | |
| BOS | 3 |
| BWI | 3 |
| DCA | 3 |
| DEN | 3 |
| DFW | 3 |
| DTW | 3 |
| FLL | 3 |
| IAD | 3 |
| IAH | 3 |
| MDW | 3 |
| MIA | 3 |
| MSP | 3 |
| PDX | 3 |
| PHX | 3 |
| SAN | 3 |
| SEA | 3 |
| | |
| CLE | 4 |

| | |
|---|---|
| CVG | 4 |
| HNL | 4 |
| MCO | 4 |
| MEM | 4 |
| PIT | 4 |
| SLC | 4 |
| STL | 4 |
| TPA | 4 |