Project Report
ATC-394

Tower Flight Data Manager Benefits Assessment: Initial Investment Decision Interim Report

T.G. Reynolds
A.L. Alexander Horrey
M.A. Ishutkina
D. Joachim
R. Jordan
A. Nakahara
A.L. Weaver

19 November 2012

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Federal Aviation Administration,
Washington, D.C. 20591

This document is available to the public through
the National Technical Information Service,
Springfield, Virginia 22161

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. ATC-394	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Tower Flight Data Manager Benefits Assessment: Initial Investment Decision Interim Report		5. Report Date 19 November 2012 6. Performing Organization Code
7. Author(s) Tom Reynolds, Amy Alexander Horrey, Mariya Ishutkina, Dale Joachim, Richard Jordan, Alex Nakahara, and Alison Weaver		8. Performing Organization Report No. ATC-394
9. Performing Organization Name and Address MIT Lincoln Laboratory 244 Wood Street Lexington, MA 02420-9108		10. Work Unit No. (TRAIS) 11. Contract or Grant No. FA8721-05-C-0002
12. Sponsoring Agency Name and Address Department of Transportation Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered Project Report 14. Sponsoring Agency Code

15. Supplementary Notes

This report is based on studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology, under Air Force Contract FA8721-05-C-0002.

16. Abstract

This document provides an overview of MIT Lincoln Laboratory's activities in support of the interim stage of the Initial Investment Decision benefits assessment for the Tower Flight Data Manager. It outlines the rationale for the focus areas, and the background, methodology, and scope in the focus areas of departure metering, sequence optimization, airport configuration optimization, and safety assessment. Estimates of the potential benefits enabled by TFDM deployment are presented for each of these areas for a subset of airports and conditions considered within the scope of the analyses. These benefits are monetized where possible. Recommendations for follow-on work, for example, to support future benefits assessment efforts for TFDM, are also discussed.

17. Key Words	18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	102	

FORM DOT F 1700.7 (8-72) Reproduction of completed page authorized

This page intentionally left blank.

REVISION HISTORY

Version	Description	Release Date
0.1	Initial working document	Not released
0.2	Initial MIT LL internal review document	1/12/2012
0.3	Initial MCR review document (not for further distribution)	1/18/2012
0.4	MIT LL internal review with updates and MCR comments	2/02/2012
1.0	Initial FAA release version	2/10/2012
1.1	Update incorporating revisions and addressing comments	3/9/2012

This page intentionally left blank.

EXECUTIVE SUMMARY

The Tower Flight Data Manager (TFDM) is an advanced tower automation system being developed to meet the Federal Aviation Administration (FAA) Next Generation Air Transportation System (NextGen) objectives. The TFDM system provides controllers 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 will also be provided with a suite of Decision Support Tools (DSTs) that provide critical information for assistance in tactical and strategic decision-making. In addition, TFDM will facilitate data exchange between controllers within a tower facility, between Air Traffic Control facilities, and between stakeholders. The capabilities provided by the TFDM system 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.

This document summarizes the analysis efforts undertaken by MIT Lincoln Laboratory to estimate the benefits from the key TFDM capabilities over the 2015–2035 timeframe in support of the Initial Investment Decision (IID) benefits assessment process. These included analyses in the areas of a departure metering DST, sequence optimization DST, airport configuration DST and safety assessments. Key benefits metrics used in the analysis include delay reduction (both net reduction experienced by the passengers in getting airborne or to the gate faster, and “engines-on” delay reduction when delay is shifted from being absorbed during active taxi to the gate with engines-off) and fuel burn reduction from “engines-on” time reductions.

The departure metering DST aims to manage push-back processes in order to avoid excessive taxi time and surface congestion by keeping aircraft on the gate or other designated location until they can be efficiently handled. Departure metering benefits in the form of “engines-on” delay reduction were estimated at a set of eight key airports around the U.S. Current-day benefits were estimated using Aviation System Performance Metrics (ASPM) data and validated with field demonstration results when available. Given uncertainty in the future year inputs, benefits estimates are presented for three different future year scenarios based on FAA’s NextGen forecasts. Results also distinguish between the case when unlimited holding capacity is available and when gates are forecast to be a constraint. Across the eight study airports, benefits estimates (accounting for gate constraints) over the 2015–2035 time period range from 727 thousand hours taxi time and 188 million gallons of fuel with a value of $456 million for the FAA-provided input case, to 4.3 million hours taxi time and 1.1 billion gallons of fuel with a value of $2.8 billion for the MIT-modified input case which accounted for some issues identified in the original FAA data.

The sequence optimization DST aims to provide guidance to controllers to achieve an optimal sequence of departures and runway crossings to improve throughput and hence reduce both net passenger and “engines-on” delay. “Advanced” sequence optimization where the sequencing occurs at the gate/ramp area was considered alongside more "basic" sequence optimization where re-sequencing occurred at the departure end of a given runway. The former can be considered a substitute for departure metering in the longer term, while the latter is a complement to departure metering in the short to medium term such that their benefits can be considered additive. Sequence optimization benefits were estimated for three different runway/taxiway configurations representative of Dallas/Fort Worth International Airport (DFW) and La Guardia Airport (LGA). Benefits of basic sequence optimization at DFW were estimated to be 194,000 hours of delay reduction at DFW 2015–2035 and 428,000 hours for advanced sequence optimization, with estimated fuel savings of 50 million gallons and 110 million gallons (valued at $122 million and $270 million at $2.43/gallon), respectively. Sensitivity of results to key input variables of fleet mix and separation requirements were also studied. Further work is required to generalize the results to a larger set of airports.

The airport configuration DST aims to provide guidance to controllers regarding when to make changes to airport configuration in order to maximize throughput and reduce net passenger and "engines-on" delay as a function of the wind environment. Performance of the airport configuration function is estimated using the unserved demand metric before and after a DST is used. Airport configuration benefits were estimated for Boston Logan International Airport (BOS) and DFW airports. The BOS aggregate taxi time savings 2015 to 2035 are estimated to be 77,000 hours, and 60,000 hours for DFW. These translate to fuel savings of 17 million gallons ($41 million) for BOS and 15 million gallons ($34 million) at DFW.

Given that DFW was common to all three DST benefits assessments, it is instructive to compare the relative benefit of each capability at that airport. Departure metering is estimated to achieve 41,000–131,000 hours (depending on the input data chosen) of taxi time reduction over the 2015–2035 timeframe, compared to 194,000 hours for basic sequence optimization, 428,000 hours for advanced sequence optimization and 60,000 hours from airport configuration optimization. While these analyses examined the impacts of the DSTs in isolation, in reality they can be interdependent and future work is needed to explore these relationships in detail. It is also important to note that departure metering is estimated to have low benefits at DFW compared to many of the other airports studied, so the fraction of total TFDM benefits attributable to departure metering is likely to be larger system-wide than these DFW numbers suggest.

The TFDM safety assessment was conducted to estimate the potential incidents and accidents that could be mitigated or prevented through the deployment of TFDM capabilities. A detailed assessment of incidents and accidents during the 2005–2009 period was conducted and the impacts of TFDM on these historical incidents was extrapolated into the future. Monetization of the reduction of personal injury and aircraft damage costs was estimated using standard FAA techniques. Monetized safety benefits of TFDM for the 2015–2035 time period were estimated to be in the range $124 million–$2.6 billion depending on whether historical accident rates were limited to the 43 Airport Surface Detection Equipment, Model X (ASDE-X) analysis airports (lower bound) or National Airspace System-wide (NAS) (upper bound).

The large range is due to the estimated prevention of at least one fatal accident every five years through TFDM deployment in the upper bound, which is not present in the lower bound.

Each of the three DST capability areas and safety assessment included in this analysis could be expanded to support further benefits assessment processes and/or additional TFDM capabilities could also be explored. In addition, MIT LL strongly recommends human-in-the-loop (HITL) studies be pursued to support the assessment of core TFDM functionality. HITLs allow assessment of phased implementation of TFDM under future traffic levels, in controlled environments, under off-nominal conditions and to examine controller acceptance (e.g., workload, situation awareness, trust). HITLs can also support TFDM assessment across a wide range of objective measures (e.g., performance, environmental, operational error) to help validate and calibrate the results from the computer modeling activities or to fill benefits gaps.

This page intentionally left blank.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Page

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS
(Continued)

LIST OF ILLUSTRATIONS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

1. INTRODUCTION

Improvements to current air traffic management technologies and techniques are required to move toward the next generation air transportation system (NextGen). Over the next several decades, the Federal Aviation Administration (FAA) projects a significant increase in air traffic in the National Airspace System (NAS). Existing air traffic control towers will need to manage this growth while meeting NextGen targets for improved efficiency of surface operations. The Tower Flight Data Manager (TFDM) is being prototyped for the FAA to help address these needs. TFDM provides controllers 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 will also be provided with a suite of Decision Support Tools (DSTs) that provide critical information for assistance in tactical and strategic decision-making. In addition, TFDM will facilitate data exchange between controllers within a tower facility, between Air Traffic Control (ATC) facilities, and between stakeholders such as airlines.

The capabilities provided by the TFDM system 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. In order to assess the viability of TFDM for wide NAS deployment, it is necessary to undertake a cost-benefit analysis to support the acquisition decision process. This includes estimating the likely costs of deployment at appropriate locations relative to the potential benefits this deployment will bring. FAA has tasked MCR to perform the formal cost-benefit analysis for phase 1 of the TFDM program (TFDM-1), covering the various key stages in the investment analysis process: the Investment Analysis Readiness Decision (IARD), the Initial Investment Decision (IID) and the Final Investment Decision (FID). MIT LL is supporting MCR in these activities by providing benefits assessment data on key TFDM capabilities. The generation of such data is highly complementary to the prototyping effort. For example, the process by which benefits are identified for TFDM 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 TFDM should possess in order to address them, and therefore, helps guide priorities for the prototype system.

MIT LL is advocating a portfolio of benefits assessment activities, including operational field testing, human in the loop simulations and computer modeling. Each strategy has different roles but together they allow a quantification of benefits associated with key metrics, such as operational performance, environmental performance, user acceptance and safety. MIT LL has published reports on the field testing activities at DFW which highlighted benefits opportunities seen in an operational setting [MIT LL (2010), MIT LL (2011)]. MIT LL's focus for the interim IID benefits assessment was directed to be in the areas of computer modeling and safety assessment of TFDM capabilities, and the findings from activities in these areas are reported in this document.

This page intentionally left blank.

2. MIT LL ANALYSIS FOCUS AREA RATIONALE

The TFDM Shortfall Report [FAA (2010)] was released in July 2010 to support the IARD stage of the investment analysis cycle. It identified a pool of monetized surface efficiency improvement potential across 44 TFDM-1 analysis airports (see Appendix A) over a 20 year lifecycle (FY2016–2035), distributed among specific inefficiency areas identified in Figure 1.

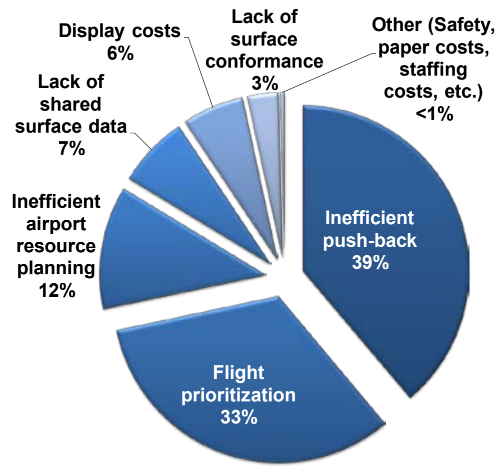


Figure 1: Shortfall Report Identified Inefficiency Breakdown

Inefficient push-back, flight prioritization and airport resource planning account for 84% of the total identified inefficiency. The first of these, inefficient push-back processes, pertains to the push-back rate exceeding the achievable departure rate, thus resulting in excessive congestion on the surface. This problem can be mitigated through departure metering procedures, which are designed to better match departure demand and capacity on the airport surface. Flight prioritization can be inefficient because airlines lack the ability to re-order their flights within a departure queue. In addition, ATC may not efficiently sequence departing flights to maximize use of airport resources. Sequence optimization decision support tools would help address some of this inefficiency. Finally, inefficient airport resource planning can arise when airports are in configurations with inferior performance relative to another available one. This inefficiency could be addressed through airport configuration optimization tools.

There is strong correlation between the tools required to address these three major inefficiency areas identified in the Shortfall Report and the capabilities recommended in the Surface Trajectory-Based Operations (STBO) mid-term concept of operations [MITRE (2010)], which have played an important role in establishing initial requirements for TFDM decision support tools. Accordingly, the focus for the

TFDM interim IID benefits assessment is placed on computer modeling of departure metering, sequence optimization and airport configuration optimization capabilities. The key benefits metrics used in the analyses include delay reduction (both net reduction experienced by the passengers in getting airborne or to the gate faster, and "engines-on" delay reduction when delay is shifted from being absorbed during active taxi to the gate with engines-off) and fuel burn reduction from "engines-on" time reductions.

Safety was seen to be a relatively small contributor to the *monetized* impacts in the TFDM Shortfall Report thanks to the remarkable safety record of the air transportation system. However, continued vigilance regarding safety impacts of changes to the ATC system is essential to maintain the high safety standards. For this reason, and because it was identified as a gap by MCR for the TFDM IID needs, MIT LL also conducted a safety assessment.

Details on the findings from activities in each of these areas are provided in the following sections in terms of background to the work, the methodologies used and key findings, followed by a summary and recommendations for follow-on work to support the Final Investment Decision process.

3. DEPARTURE METERING

3.1 BACKGROUND

Queues of aircraft at the end of departure runways ensure that there is a constant supply of aircraft for controllers to select for release and, hence, make full use of departure runway capacity. However, only a certain number of departing aircraft are required in the queue to ensure that high throughput is achieved (with this number depending on airport and operations-specific variables). When this number is exceeded, there is unnecessary additional congestion on the surface, leading to increased taxi-out delay and excessive fuel burn and emissions without any operational advantage.

One mechanism for regulating how many aircraft enter a departure queue is to throttle the number of aircraft pushing-back or leaving the non-movement ramp area in a given time period. Departure metering processes attempt to manage push-back operations and, thereby, avoid releasing aircraft onto the taxiways until they can be efficiently accommodated at the runways, while simultaneously maintaining high departure throughput. This effectively shifts taxi-out delay to the gate or ramp areas (preferably with engines off), resulting in benefits such as reduced fuel burn and emissions, increased passenger and bag connectivity, and more predictable taxi-out times.

Several implementation options to achieve the principles described above have evolved: aggregate traffic metering using a pushback rate, metering groups of flights, and metering individual flights:

- Aggregate traffic metering using pushback rate: recommend a push-back rate to ATC, who then determine which flights to clear to push consistent with this rate. This approach has the benefit of achieving some level of departure queue management with minimal real-time airline coordination and automation needs, but with lower levels of control as a result. An approach falling into this category is the “Pushback Rate Control” concept being explored by MIT with recent field trials at Boston Logan Airport [Simaiakis et al. (2011)].

- Metering groups of flights: recommend a maximum number of flights to release from a given ramp area in a certain time interval. This approach is attractive because it provides an equitable allocation of departure capacity to each flight operator who are then empowered to choose which specific flights to release given internal priorities. An approach falling into this category has been developed, successfully evaluated, and continues to be refined by FAA/STBO (Collaborative Departure Queue Management (CDQM) [Brinton et al. (2011)], demonstrated at Memphis and Orlando airports).

- Metering individual flights: recommend when specific flights should leave from gate or spot. This approach is attractive because in theory departure metering can be combined with efficient departure sequence generation in order to maximize benefits potential. However, it requires significant real-time airline coordination to know when flights will be ready to push back, as well as effective implementation of an arbitration strategy when the number ready to push exceeds the number that should be allowed to push. Approaches falling into this category are being developed by NASA Ames (Spot and Runway Departure Advisory (SARDA) [Hoang et al. (2011)]); FAA/STBO (Collaborative Departure Scheduling (CDS) [Brinton et al. (2011)]) and PASSUR Aerospace (who have been conducting field trials recently at New York John F. Kennedy International Airport JFK) [Nakahara et al. (2011)].

3.2 ANALYSIS METHODOLOGY

The high-level methodology for the benefits assessment of the departure metering capability is illustrated in Figure 2. The top half of the methodology identifies the analysis for current year benefits estimation, while the bottom half is the equivalent for future year analysis.

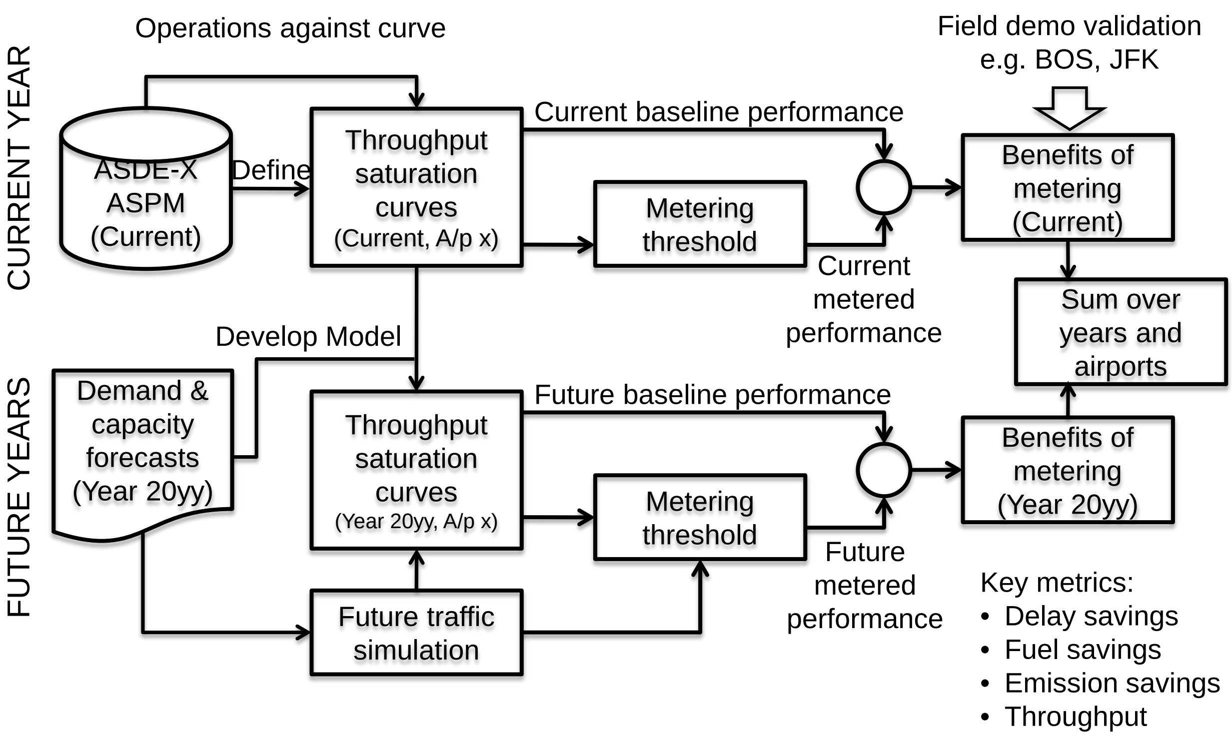


Figure 2: Departure Metering Analysis Methodology

3.2.1 Current Year Analysis

At the core of the methodology is the concept of throughput saturation curves illustrated in Figure 3, which relate departure throughput to an appropriate traffic metric (e.g., number of departing aircraft on the airport surface or in a departure queue). All of the implementation approaches described above can be characterized by this type of curve. As more aircraft push back from their gates onto the taxiway system, the throughput of the departure runway initially increases because more aircraft are available in the departure queue. But as the number of aircraft continues to increase, the airport eventually reaches a saturation departure rate. The saturation value depends on the airport configuration, arrival demand and meteorological conditions (VMC/IMC) as well as other airport conditions and human factors considerations. Beyond the saturation point, any additional aircraft that push back simply increase the time they are taxiing out with engines on without any gain in departure rate. The objective of departure metering is to maintain the number of aircraft pushed back at (or below) a certain control level just above the saturation point. In this way, high departure throughput can be achieved without unnecessary surface congestion and the resulting excess delays and fuel burn.

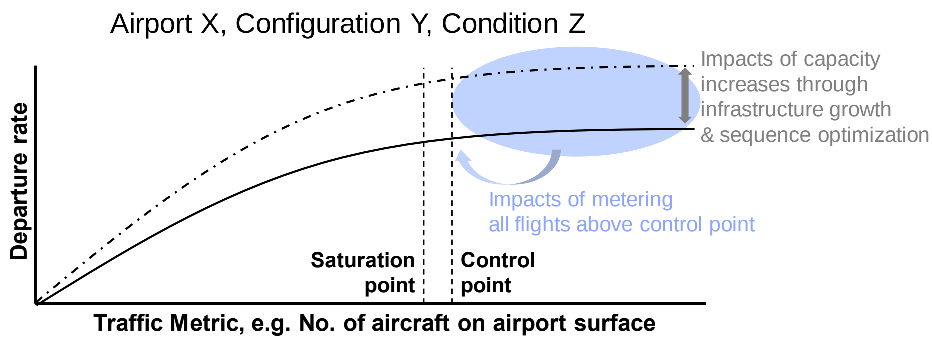


Figure 3: Throughput Saturation Curve

The inputs identified in Figure 2 allow current-day throughput saturation curves to be determined for study airports using archived operational data such as ASPM [FAA/ASPM (2011)] and ASDE-X surveillance. ASPM data covering the 2010 period was used to characterize current day operations, with the exception of JFK airport (where 2009 data was used) due to the runway construction and departure metering field trials which took place in 2010. ASPM data provides OOOI (OUT, OFF, ON, IN) times for individual flights, as well as airport configuration in 15 minute intervals, while the OFF times can be used to calculate the airport throughput in the same 15 minute intervals. From this, the appropriate traffic metrics can be related to departure throughput as function of configuration, arrival demand and weather conditions. Once the throughput saturation curves have been determined, the operational data also allows a determination of where the airport is operating upon that curve at different times. It is then possible to determine the aggregate taxi delay due to congestion for different time intervals or operating conditions.

The average taxi time of flights taxiing when the number of actively taxiing aircraft, *N* is at the control value (*N_Control*) is compared to the average taxi time of flights taxiing when *N* > *N_Control*. The number of flights in the congested regime is also calculated and the potential benefits in terms of taxi time reduction are then calculated as:

$$\text{Active Taxi Time Reduction Benefits} = \#\ \text{Flights}_{\text{Congestion}} \times (\text{AverageTaxiTime}_{\text{Congestion}} - \text{AverageTaxiTime}_{\text{Control}})$$

Using ICAO-standard fuel flow and emissions rates for different engine types [ICAO (2011)] and matching these to aircraft types used at study airports, the taxi time can be converted to fuel burn and emissions under the operating conditions observed.

Note the location of the control point relative to the saturation point is a trade-off between risk of reducing the departure rate (control points much higher than the saturation point reduce this risk) versus achieved benefit (control points closer to the saturation point increase the benefit): this trade-off is explored in the analysis to better understand the range of potential benefits from metering capabilities as a function of the control point. Control points closer to the saturation point increase the benefit obtained because taxi time scales roughly linearly with the traffic metric but can reduce throughput and increase delay if they are too far below the saturation point. The curve in Figure 4 shows taxi time benefits values decreasing as *N_Control* increases, which is typical across airports. This occurs because there are fewer and fewer aircraft affected as *N_Control* grows, resulting in benefits that slowly taper towards 0. At some airports with higher average *N* values, the decrease in benefits is more steady as *N_Control* increases. The typical value for *N_Control* at JFK is between 25 and 30. Since these curves are similar for all airports, the other airports are presented in Appendix B. The values for *N_Control* used in this study are equal to the value predicted by the Random Forest (RF) model for the saturation point. The saturation point is defined in this study as the first point at which the throughput reaches 95% of its maximum value. To eliminate the high variability due to small sample sizes, the 2% of data with the highest *N* values is removed from the data set.

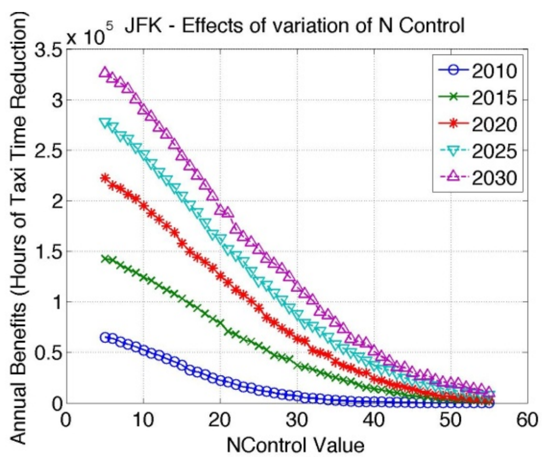


Figure 4: Variation of Metering Benefit with N_Control Value

Achievable benefits are also constrained at some airports by factors such as the number of available gates and holding space. Therefore, these impacts are explicitly modeled for the analysis airports. The predicted metering benefits using the approach outlined above with current day throughput saturation curves are validated against benefits data from operational field demonstrations at JFK [Nakahara et al. (2011)] and BOS [Simaiakis et al. (2011)] to provide confidence in the realism of the outputs from the TFDM analysis approach.

3.2.2 Future Year Analysis

Prior work at MIT has automated the determination of benefits for current day operations using the approach outlined above [Simaiakis (2009)]. This approach has been expanded for the TFDM IID benefits assessment application to explore future year impacts. The key sources of future year data for use in this (and the other) analysis include:

- Forecast annual demand by airport: FAA Terminal Area Forecast (TAF) 2010–2030 [FAA/TAF (2011)].

- Forecast capacity by airport: MITRE FACT2 [FAA/MITRE (2007)] and 2011 update [personal communication, Tom Reynolds (MIT LL)/Dan Howell (MCR), 3/21/2011].

- Future year "NextGen schedules" for 12 representative days in the years 2010, 2015, 2020, 2025 and 2030 [personal communication, Tom Reynolds (MIT LL)/Kimberly Noonan (FAA/AJP), 4/28/2011 with revisions provided by MCR, 12/7/2011]. These include untrimmed, trimmed and "fleet evolved" schedules.

- Physical airport data (Terminals, gates, runways, etc.) from airport websites.

Future year benefits were calculated by simulating saturation curves and congestion at each study airport for the future "out-years" of 2015, 2020, 2025, 2030 and 2035. This required the development of a three-stage model which calculated the future saturation curves, calculated the operating point on those curves, and calculated the benefits from metering, as illustrated in Figure 5. Given uncertainty in the future year inputs, benefits estimates are presented for three different future year scenarios based on FAA's NextGen forecasts termed "FAA Original," "FAA Adjusted," and "MIT Simulation."

The first were the original future year NextGen schedules provided by the FAA from their System Wide Analysis Capability (SWAC) model [Noonan (2011)]. This is a NAS-wide network model which is used by the FAA to develop NextGen flight scenarios, and has the ability to "trim" flights when the capacity:demand ratio at any given node (airport) in the network causes degraded performance beyond a certain threshold level. Figure 5 (FAA Original Schedule, black path) shows that it is not necessary to use the MIT models and simulations for this input path because the schedule provided by the SWAC model contains Gate Out and Wheels Off times from which saturation throughput curves can be derived. In the course of examining the schedules, however, two major concerns were raised which led to the use of two other inputs for the model (FAA Adjusted and MIT Simulation) to generate comparative results, as described below.

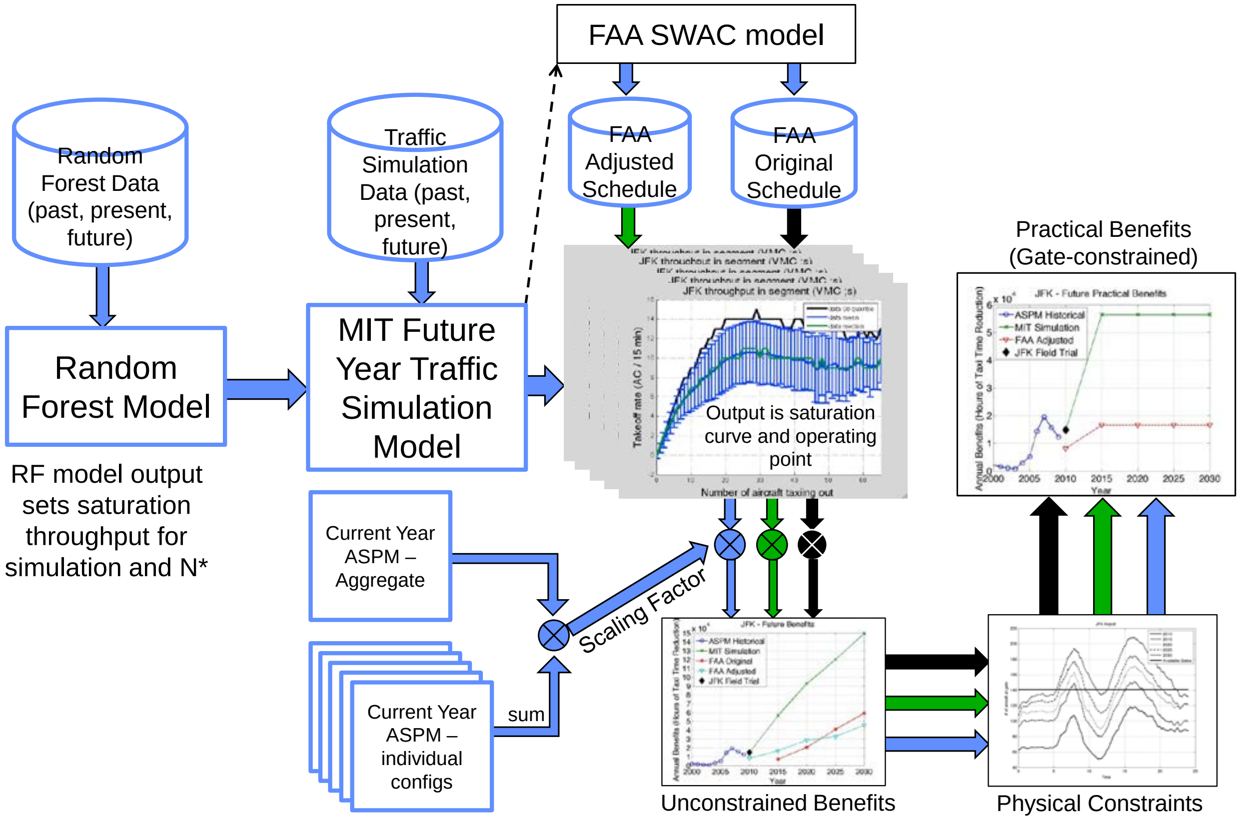


Figure 5: Future Year Benefits Assessment Detailed Methodology

The first identified problem with the FAA Original data is that the capacities used to trim the schedules are very optimistic. They represent the 'peak' capacity at the airport, but due to the increase in traffic predicted at most airports, the capacities end up being reached consistently in the schedules. For instance, in 2030 at JFK the airport operates at 60 departures an hour between the hours of 3 PM and 9 PM. Assuming one departure runway (common at JFK), this equates to one departure/minute. This rate is seen occasionally in the past, but never for six consecutive hours. From queuing theory, the expected delay increases nonlinearly as demand approaches capacity, which is why airports cannot operate at their capacity for an extended period. In addition, the fleet mix at JFK would require separations of more than 1 minute, on average, for wake vortex mitigation since there are many large and heavy aircraft. The conclusion from this is that while the peak capacity of JFK might be 60 movements/hour, the sustained capacity (which is more applicable in this case) is probably between 40 and 50 movements/hour. If the peak capacity is used, we are concerned the forecast demand cannot be met with the existing facilities at JFK.

To address this first concern, the capacities used to calculate the FAA Original schedules were scaled down to produce the “FAA Adjusted” future year schedule. To derive the scaling factors, the average saturation departure throughput across the five most common configurations observed in the ASPM data for 2010 was compared to the FAA Original implied capacities at each study airport for 2010. Ten percent was added to this ratio at most airports to account for variability in the saturation throughput. At DFW and Washington Dulles International Airport (IAD), the scaling factor was set to 1 to reflect the fact that those airports never truly reach saturation (i.e., there is not enough demand to warrant using their full capacity). Table 1 shows the FAA Original capacities (first column) as well as the average saturation throughput, and the scaling factor used to produce the FAA Adjusted inputs. For the capacities for 2015 and beyond, the FAA Adjusted values grew at the same percentage as the FAA Original values. The SWAC model was then re-run with the scaled down capacities to produce the FAA Adjusted inputs, resulting in less demand at some airports due to the trimming algorithm. The benefits calculated with the new schedules are the FAA Adjusted inputs (Green path in Figure 5).

Table 1: Scaling Factors to Derive FAA Adjusted Future Year Schedules

	2010 Hourly Departure Capacities			
Airport	2010 FAA Original	2010 ASPM Observed Saturation	FAA Original/ASPM Ratio	FAA Adjusted Scaling Factor
ATL	149.7	113.7	0.76	0.85
BOS	82.9	40.2	0.48	0.6
DFW	125.8	N/A	N/A	1.0
IAD	96.7	N/A	N/A	1.0
JFK	59	41.8	0.71	0.8
LGA	48.5	35	0.72	0.75
ORD	124.3	94.1	0.76	0.85
PHL	108.9	45.6	0.42	0.5

The second concern is the simulation used to predict taxi times. Once again examining JFK, Figure 6 shows the distribution of taxi times in the FAA Original SWAC schedule for 2010, the actual taxi times in 2009 from ASPM and the MIT simulated taxi times for 2010. The FAA Original schedule underpredicts large taxi times ($T > 50$) and this is also seen when the taxi times are estimated into the future, as shown in Figure 7.

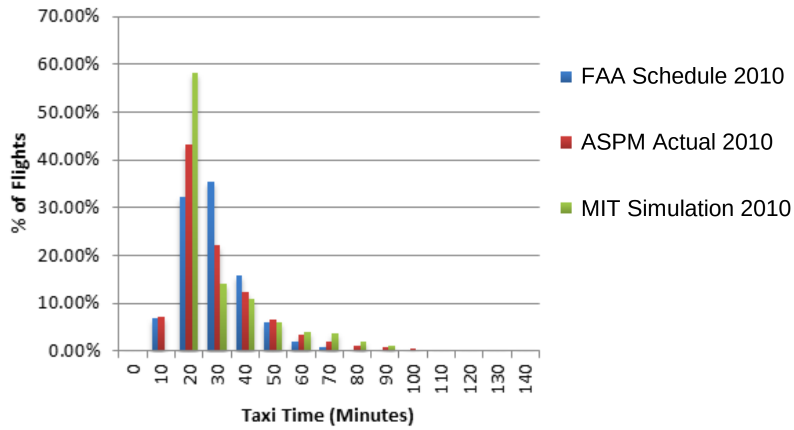


Figure 6: JFK 2010 Average Taxi Time Comparison

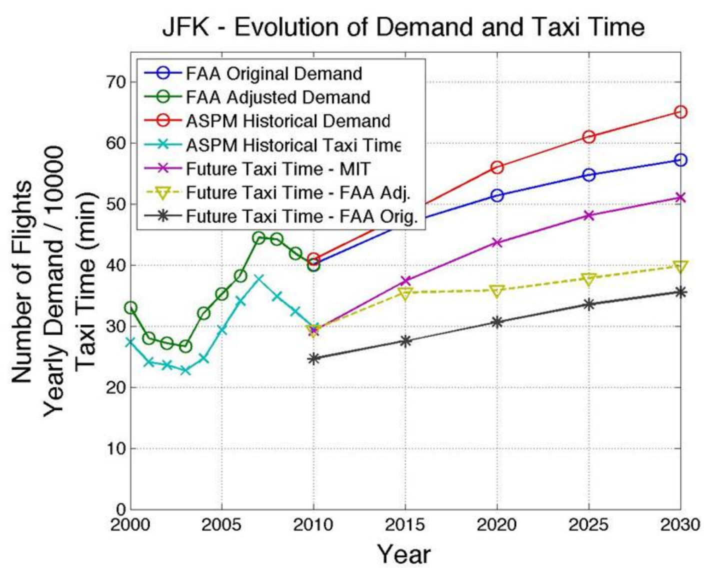


Figure 7: JFK Future Year Average Taxi Time Comparison

The FAA Adjusted schedules still do not address the second concern, the taxi time calculation, which is why the MIT Simulation was developed. While the MIT distribution does not perfectly match the ASPM data in Figure 6, there is significant variation by configuration. The MIT Simulation matches the ASPM distribution much better for large taxi times; the FAA model has almost no flights with taxi times of more than 60 minutes. Since this study examines congestion, which is when these large taxi times occur, it is important to get a better model for long taxi times. Figure 7 shows that the average taxi times are actually well predicted by both FAA Adjusted and MIT Simulation cases in 2010. The demand in 2015 is slightly (4%) larger in 2015 than in 2007, the previous peak in taxi times. However, the FAA Adjusted model does not reach 2007 levels until 2025 when the demand is 20% higher than 2007, while the MIT model matches the 2007 average in 2015 as would be expected. The figure also shows that taxi times are strongly related to the demand level, as seen in the historical data. The future FAA taxi times do not show the same relationship (unlike the MIT estimate), which is further validation of the MIT modeling approach.

In the first stage of the MIT Simulation, the future performance of the airport was estimated using the Random Forest (RF) method [Breiman (2001)]. This method was chosen because of the many parameters and conditions that affect an airport's performance, as well as the nonlinearity of the performance. A simple regression is unable to accurately capture all these behaviors. 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 departure metering because it makes no assumptions about the functional relationship between the input/predictor variables and the output, avoiding a bias that would be introduced by assuming a particular function is the correct form to describe airport behavior. In the departure metering analysis case, the saturation point and saturation throughput are the target prediction variables which define the airport throughput saturation curves to first order. The input parameters to the random forest model were chosen using engineering judgement as well as the input of experts at MIT LL, FAA and MCR and included the mean and peak hourly demand and capacity, the frequency of usage of a configuration, the physical size of the airport, and the number of gates. Appendix B contains a full list of the variables as well as their source. The decision trees were trained, or 'grown,' on data from 2000 to 2010, those being the years for which appropriate 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 curve parameter estimates.

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 [Simaiakis & Pyrgiotis (2010)] was modified to use the inputs identified from the RF method to predict taxi times in the future. The simulation calculates the time for three different segments of taxiing: unimpeded time, taxiway congestion time, and time in departure queue. These three segments have tunable parameters: the average unimpeded time (by airline or overall), the taxiway congestion factor, and the capacity. The average unimpeded time used was the overall average from 2010, since the physical layout of an airport can change suddenly and unexpectedly, e.g., if an airline relocates terminals. By using the overall taxi time, the robustness of the model is improved. The taxiway congestion factor is calibrated from the present day training data. The capacity is tuned using the output from the Random Forest model (saturation throughput). Benefits of metering relative to the baseline case for future years were calculated in the same way as described above for the current-day case.

Due to the nature of the traffic simulation, the benefits are calculated by configuration. This means that the saturation curve and congestion are calculated multiple times per airport, once for each configuration that is in use. The 2010 reported weather and configuration are copied and used in 2015, 2020, 2025, and 2030. This assumes that the weather and configuration choices in that year are representative of the future. A possibility for future work would be to combine the airport configuration optimization and departure metering studies to explore the possible added benefits from using the optimal configuration at all times. The reason behind calculating benefits by configuration is that elements such as unimpeded time and capacity in the traffic simulation are very dependent on the configuration. Training and testing the simulation on aggregate, airport-wide results greatly degraded the accuracy of the simulation. In addition, the chosen approach identifies configurations and times where metering is not needed or beneficial. For example, at Boston Logan the '4' configurations show little benefit from metering relative to the others. This matches observations seen in the MIT Field Trial [Simaiakis et al. (2011)].

To calculate the full year benefits, the results from the individual configuration were summed and scaled up. Because only a few configurations have enough data to train the traffic simulation accurately, the five most common configurations in the base year were chosen to be simulated at each study airport. In addition, only VMC conditions were looked at because IMC conditions also happened too rarely to train the simulation. The benefits from each of the five configurations were summed and scaled by the percentage of departures a given configuration handled out of the yearly total. This choice caused an inconsistency between the simulation results and the field trial results for two reasons: firstly, low-use configurations can have disproportionate benefits from surface congestion management, meaning that only scaling the top five configurations underestimates the true benefits. Secondly, only examining VMC conditions also underestimates the benefits since IMC conditions have more benefits/hour of time. To account for this, the benefits for the base year were calculated 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 output from this approach was an estimate of the "unconstrained benefits" at a given study airport. However, as previously described, in reality there are physical constraints to the number of flights that can be held by a surface congestion management approach, e.g., by the number of gates or off-gate hold locations. The "practical benefits" need to consider airport gates as a limiting resource. To predict when gate availability could become a constraint, gate utilization was assessed at the study airports. OOOI times from ASPM were used to create a running count of the number of aircraft at a gate over the course of the day. This count was calculated by incrementing a counter by one when an aircraft arrived (IN time) and decreasing the counter by one when an aircraft departed (OUT time). 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 (e.g., in the U.S. gates are 'owned' by a specific airline and are not generally 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 physical space-constrained airport in the initial study set) identified several possible locations.

The gate utilization was calculated for each airport and year in the study and compared to the number of gates at the airport (as of 2010). If the analysis showed that there would not be enough gates to accommodate unconstrained surface congestion management for more than approximately 1–2 hours during the day, then the benefits were restricted to the last year in which there were sufficient gates. The only exception to this policy is ORD, where there is space for future terminals to be constructed which would alleviate the gate usage problem. ORD benefits were still constrained, but due to excessive delay, as will be discussed in Section 3.4.7. Figure 8 shows a sample chart using historical data from LGA in 2010 with the average (Total), 95^{th} and 5^{th} percentile gate utilization. With 73 gates at LGA, there is over 100% utilization overnight, suggesting that some flights are being parked off-gate. The utilization then drops to around 50% during the day, showing there is sufficient gate space to conduct holds of over 35 flights at any given time using the simplifying assumptions outlined above.

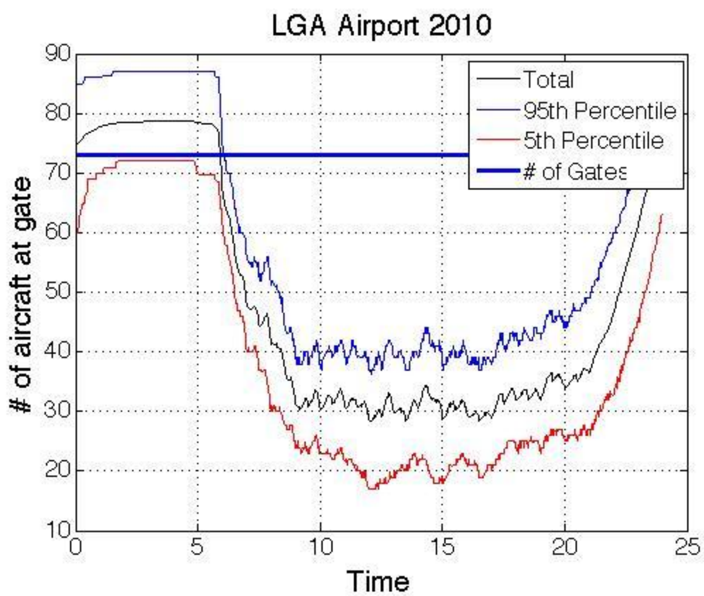


Figure 8: LGA Gate Utilization (ASPM 2010)

3.3 AIRPORT SCOPE

Eight airports have been examined in detail for the IID: ATL, BOS, DFW, IAD, JFK, LGA, ORD, and PHL. JFK and BOS were chosen because of their recent and ongoing field trials of departure metering providing important validation cases. The other airports were chosen to represent different types of airports. LGA is congested 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.

3.4 RESULTS

Results are given for the eight study airports in terms of hours of taxi time reduction. In the discussion section, these time savings are converted into fuel and monetary savings. Three unconstrained benefits curves are presented in the top right panel for each airport representing the three input cases: FAA Original, FAA Adjusted and MIT Simulation. In general, the FAA Original and FAA Adjusted are similar, but much less than the MIT Simulation results. The exceptions will be discussed in the sections for the individual airports. The reason for the large discrepancies is a derivative of Figure 6. If taxi times in congestion are being systematically underpredicted, then congestion will correspondingly be underpredicted, leading to a lower need for departure metering and hence lower benefit estimates. 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 MIT methodology because the 2010 (ASPM) and 2010 (MIT) 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 which helps to interpret the unconstrained benefits results. The difference in future demand represents the additional trimming done by MIT. The historical FAA capacity is the average declared Airport Departure Rate (ADR) from ASPM, while the MIT capacity is the saturation throughput. Note that the FAA Original capacity estimates do not correspond with the historical declared capacity at most airports. The bottom left chart for each airport shows the average gate utilization during the study years for the FAA Adjusted and MIT Simulation input cases (the FAA Original results are similar in shape but reflect higher gate usage due to the higher forecast demand in that data set), 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. Benefits are calculated from 2015 to 2035, but the 2035 point will be the same as the 2030 point since there is no input data for that year.

3.4.1 ATL Airport

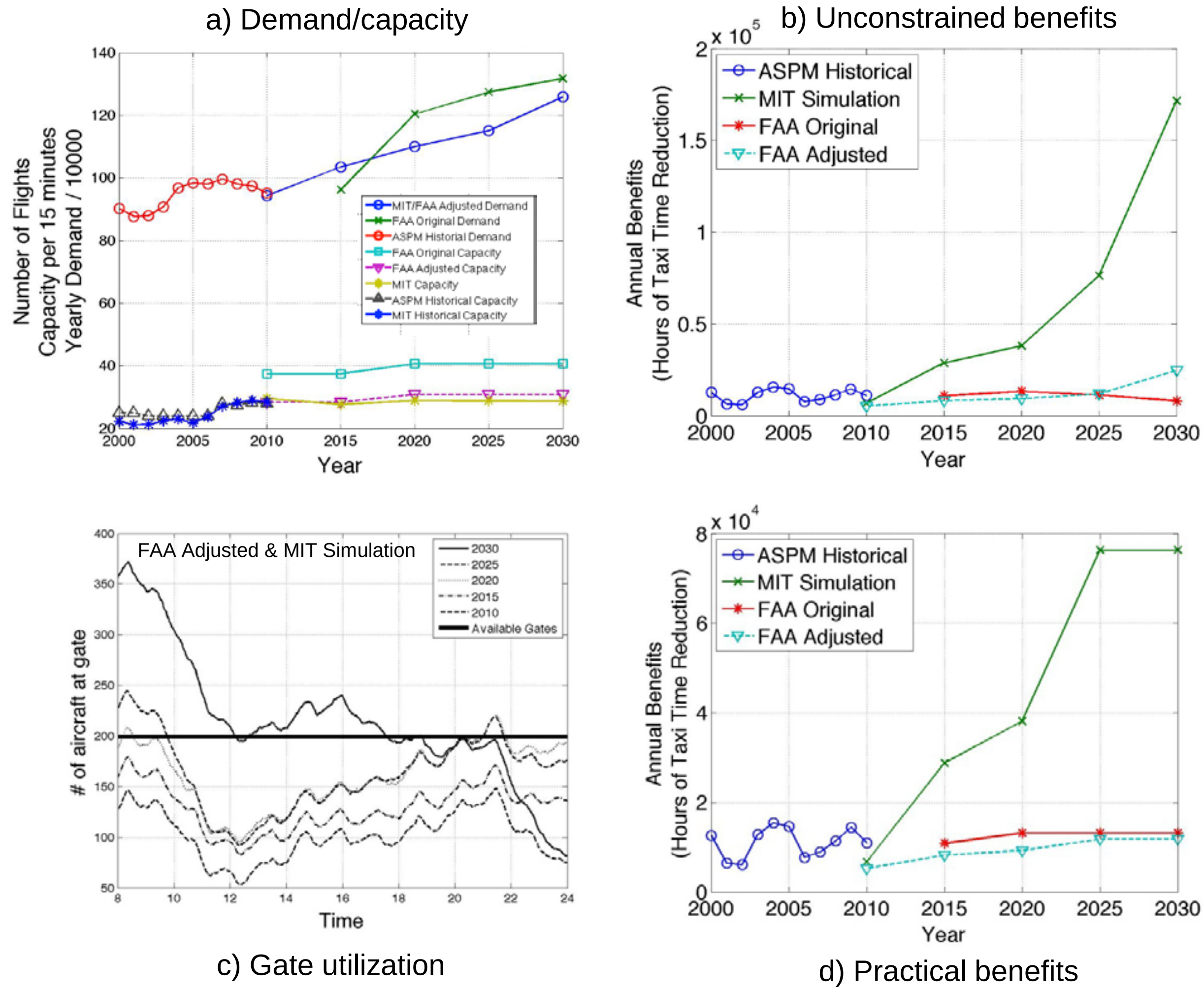


Figure 9: ATL Departure Metering Benefits Results

As shown in Figure 9, demand at ATL is predicted to increase by about 30% from 2010 to 2030, while the predicted capacity does not increase significantly. This is true for both the FAA Original and FAA Adjusted case. The lower demand in 2015 for the FAA Original compared to the Adjusted case is probably an artifact of the trimming mechanism used for the Adjusted case and has no obvious significance. The difference in capacity curves show that Atlanta's forecasted capacity is moderately higher than the historical declared capacity, and that the declared capacity and the saturation throughput are similar. The historical benefits show an increase in 2003–4 as a result of the increasing demand, before dropping off as the airport performance improved from 2006–8 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 FAA Original and Adjusted benefit curves are both relatively low because of the combination of an overestimate in capacity and an underestimate in taxi time in those data sets. 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. The early morning high gate utilization in 2030 will be explained in the discussion section. The gate utilization shows a steady scaling up of the current pattern with increasing demand. Analyzing the shape of the curve is beyond the scope of this study, 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).

3.4.2 BOS Airport

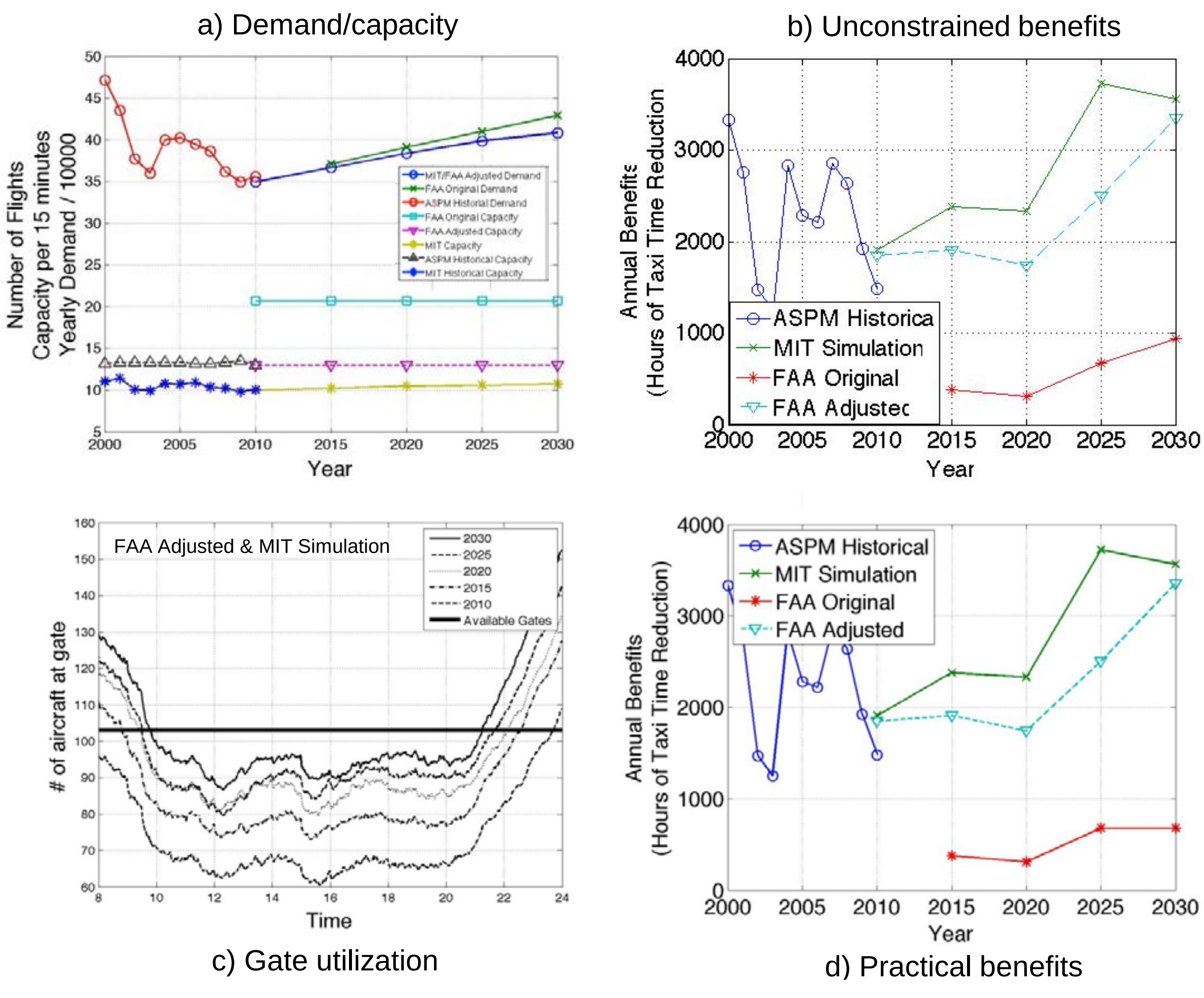


Figure 10: BOS Departure Metering Benefits Results

As shown in Figure 10, 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 forecasted capacity is significantly scaled down in the FAA Adjusted case, but it does not result in a significant decrease in demand since the airport is not forecast to reach its full capacity. The historical benefits roughly follow the historical demand at BOS. Since 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 since 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 MIT Field Trial results are shown in Figure 11 and is separate because it was only conducted on two configurations and over 1 month making a year-long estimate of benefits difficult. However, the field demonstration results compare favorably to the configuration specific benefits estimated using the MIT Simulation approach, providing some validation of the operational realism afforded by the approach.

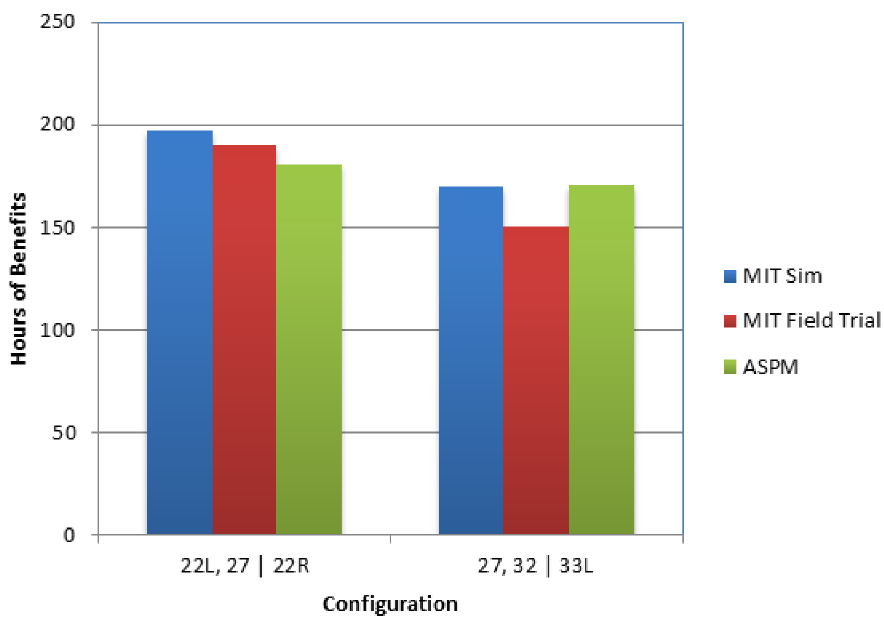


Figure 11: Boston Field Demonstration Validation

3.4.3 DFW Airport

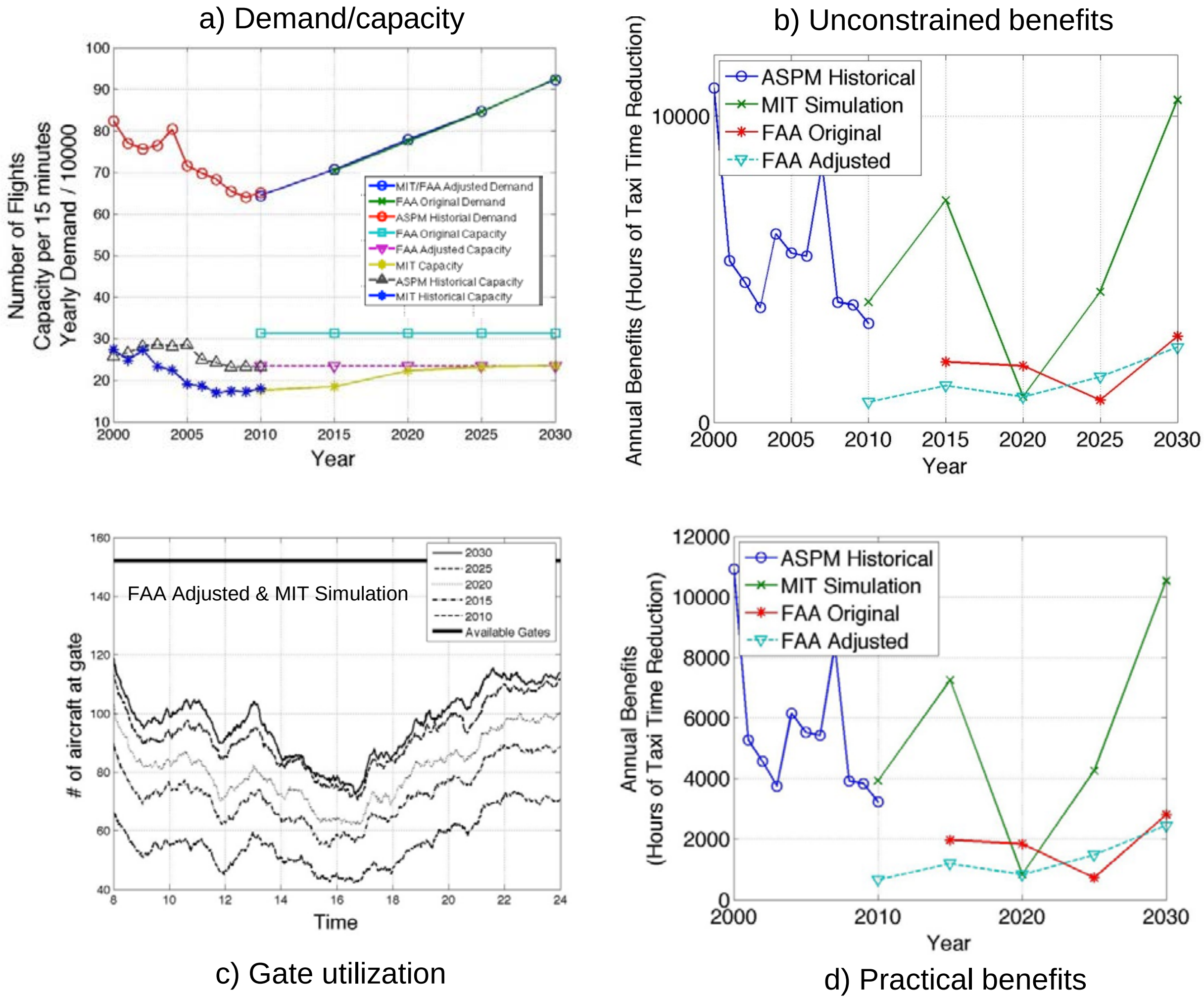


Figure 12: DFW Departure Metering Benefits Results

As shown in Figure 12, DFW airport is forecast to reverse its recent decline in traffic, and there is no expected increase in capacity. The sudden change in the MIT benefits in 2020 is due to the MIT 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 similar volatility. The reason for the drop is the gain in throughput (MIT capacity) between 2015 and 2020. This is a reversion to the performance of 2000, when there was a similar demand level. The benefit spike in 2007 is partly due to the drop in performance since 2000. This effect is elaborated on in the discussion section. 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.

3.4.4 IAD Airport

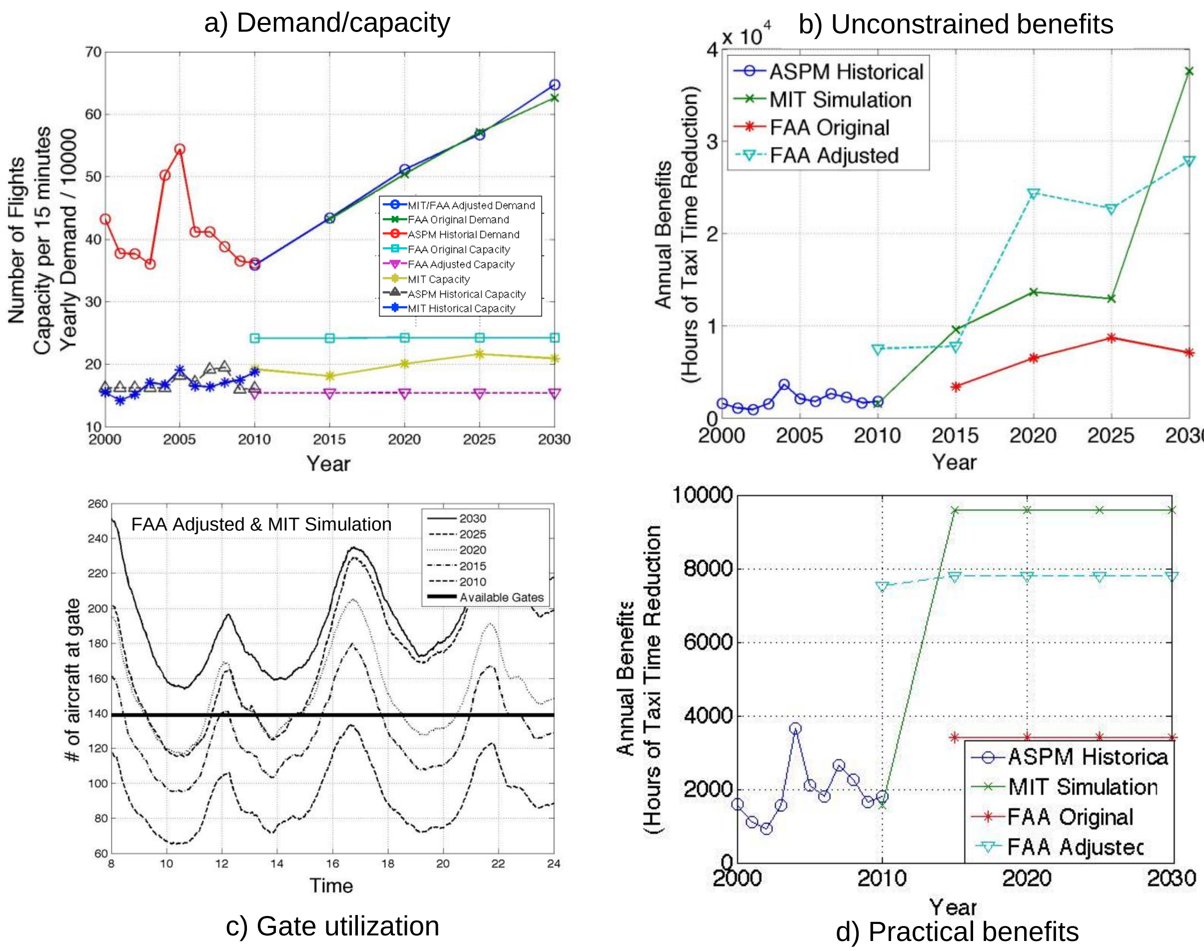


Figure 13: IAD Departure Metering Benefits Results

As shown in Figure 13, Dulles airport shows steadily increasing traffic from 2010 through 2030. There are several discrepancies in the results that could merit further analysis and all of these factors should be carefully considered when interpreting the IAD results. The first is the spike in traffic in 2004 and 2005 that did not produce the same magnitude increase in benefits relative to other airports. This could be because IAD is operating far below its available capacity, even in 2004–5. However, traffic in 2020 is predicted to be at 2005 levels and yet the predicted benefits are much higher. There is also a gap between the 2010 ASPM and FAA Adjusted benefits, although the ASPM, FAA Original and MIT Simulation estimates are equal. Even the FAA Original estimates of benefits are relatively high in future years, even though they have been shown to be very low compared to the historic benefits at other airports. A partial cause of the high future benefits is shown in Figure 14. 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 13a matches with the FAA 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 ASPM 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 to account for the higher demand levels. The IAD saturation curves from 2005 and 2006 (Figure 15) 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. As a result, the benefits at IAD are much more uncertain than at the other airports. The combination of this issue with the discrepancy in the number of flights means that the benefits at IAD have an upper bound at the MIT Simulation level and a lower bound around the FAA Original level. In terms of gate utilization, Figure 13 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.

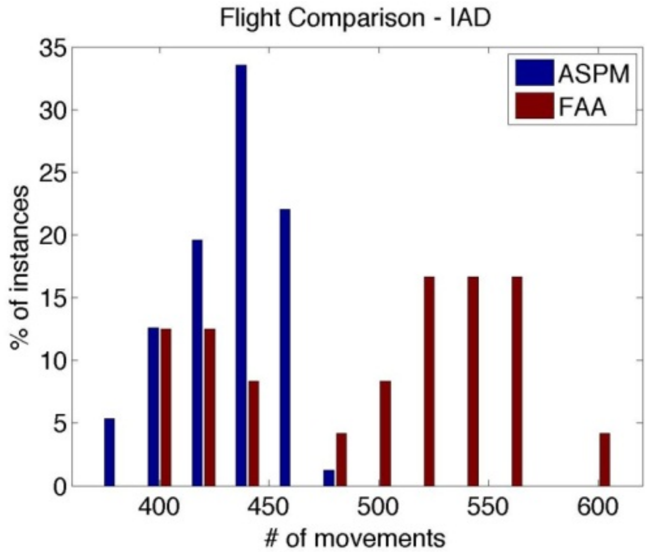


Figure 14: Average Number of Departures/Day in ASPM and FAA Schedule (2010)

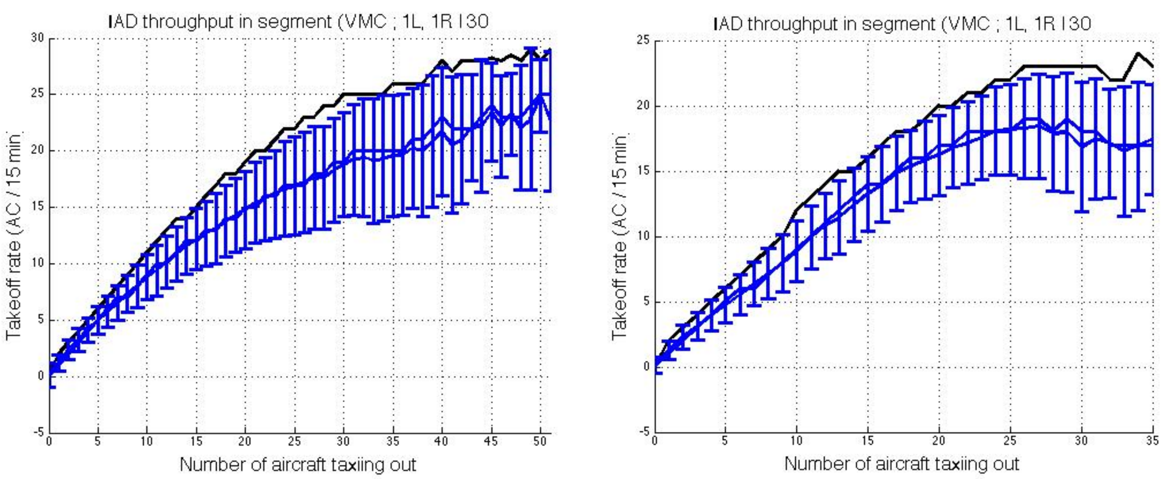


Figure 15: IAD Throughput Saturation Curves for 2005 (left) and 2006 (right)

3.4.5 JFK Airport

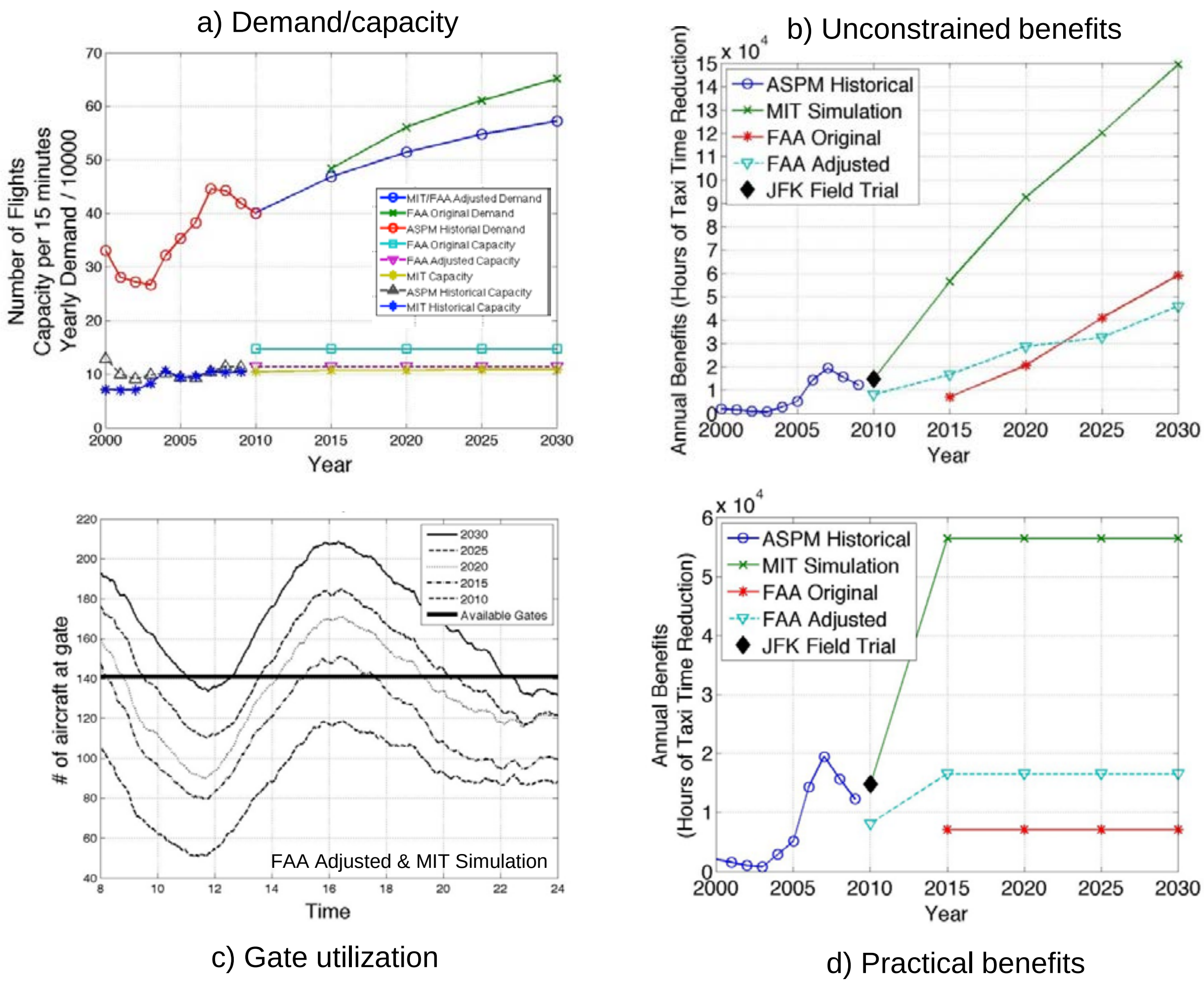


Figure 16: JFK Departure Metering Benefits Results

As shown in Figure 16, JFK shows steadily increasing benefits as the demand increases. The FAA Original demand is significantly higher than the Adjusted demand, and both are above the peak historic traffic level seen at JFK in 2007 even though the FAA capacity is not forecasted to grow. The MIT capacity is forecast to grow slightly as a response to the increasing demand. In terms of benefits, the FAA Original and Adjusted cases are both significantly below the MIT Simulation levels, which is due to the taxi time under-predictions presented in Figure 6. 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 will not be met. 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 allows further validation of the MIT approach. The JFK field trial results reported an annual taxi time reduction benefit of approximately 15,000 hours in 2010 [Nakahara et al. (2011)]. This is shown as the black diamond in panels b and d, and the comparison with the MIT Simulation results are again very favorable.

3.4.6 LGA Airport

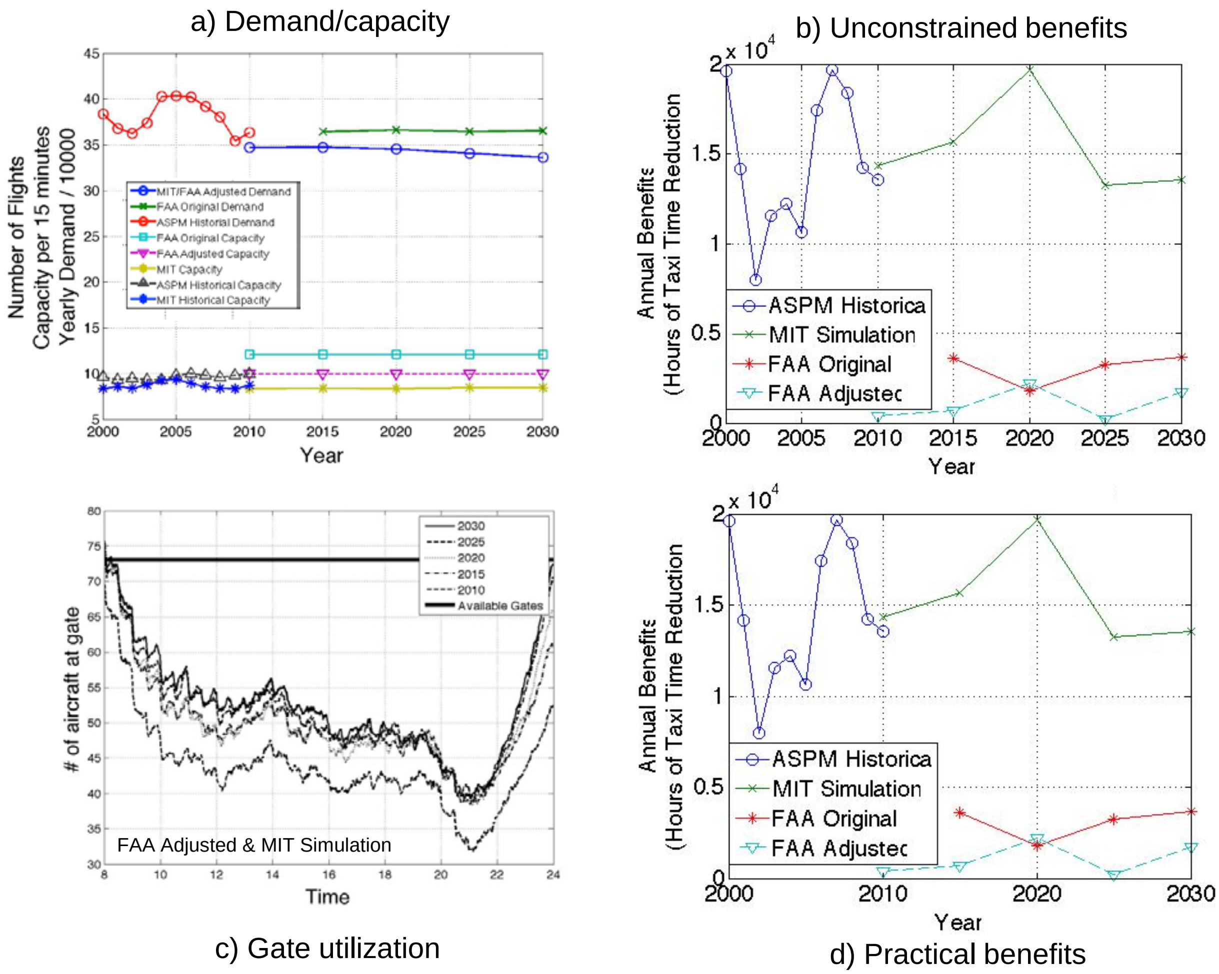


Figure 17: LGA Departure Metering Benefits Results

As shown in Figure 17, LGA is forecast to have little 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 and does not have a major impact on overall benefits. The FAA benefits are much lower because of the discrepancy seen in Figure 18. This shows the FAA-implied saturation curve in blue and the MIT simulated curve in red/green. The FAA curve does not reach the saturation throughput of 9 aircraft/15 minutes until N = 35, in contrast to the MIT Simulation (and historical data) of which reaches the saturation throughput nearer N = 15 aircraft. Since the value is so high in the FAA curve, the benefits are low because there are very few aircraft taxiing with N > 35. In addition, the underestimation of taxi times adds to the lack of 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 8, which serves as a reality check.

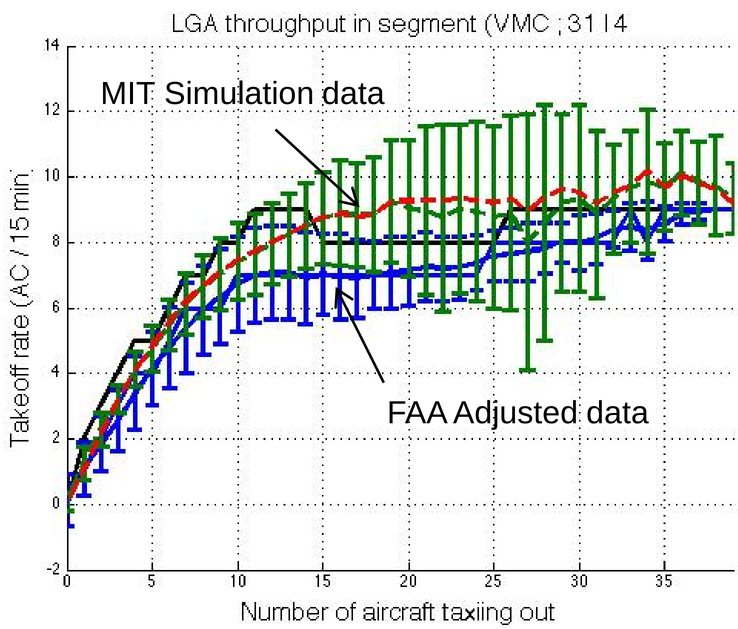


Figure 18: LGA 2030 Throughput Saturation Curve Comparison

3.4.7 ORD Airport

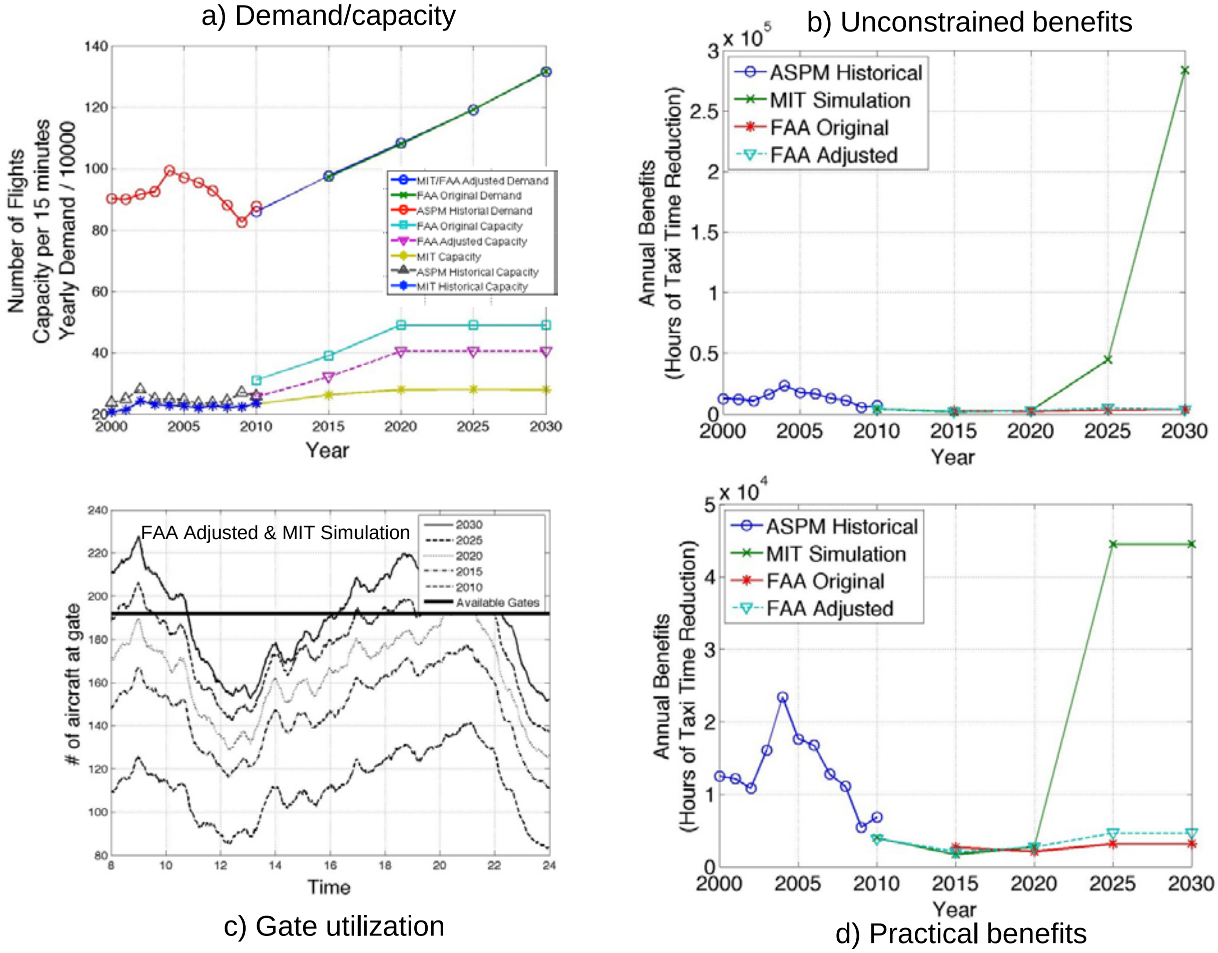


Figure 19: ORD Departure Metering Benefits Results

As shown in Figure 19, 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 major physical expansion. The expansion introduces much uncertainty into the calculation of benefits because the capacity of the airport determines the demand (through trimming) as well as the performance. The MIT Simulation implies high levels of benefits in 2030 because it predicts that the capacity of O'Hare will not be as great as the FAA models. This prediction is based on the past performance of O'Hare as well as the performance of DFW, an airport whose current configuration is similar to O'Hare's future configuration. The demand will thus exceed the capacity at O'Hare in the MIT model, causing a large amount of congestion. The future O'Hare 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 MIT 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 FAA Original or Adjusted levels, the demand would remain the same but the benefits would again be lower because there would be fewer delays (but probably not as low as the FAA benefits because of the taxi time underestimation). For these reasons, the MIT practical benefits for 2030 were capped at the 2025 level. Demand for metering can be satisfied with current terminal infrastructure until 2030. However, O'Hare 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.

3.4.8 PHL Airport

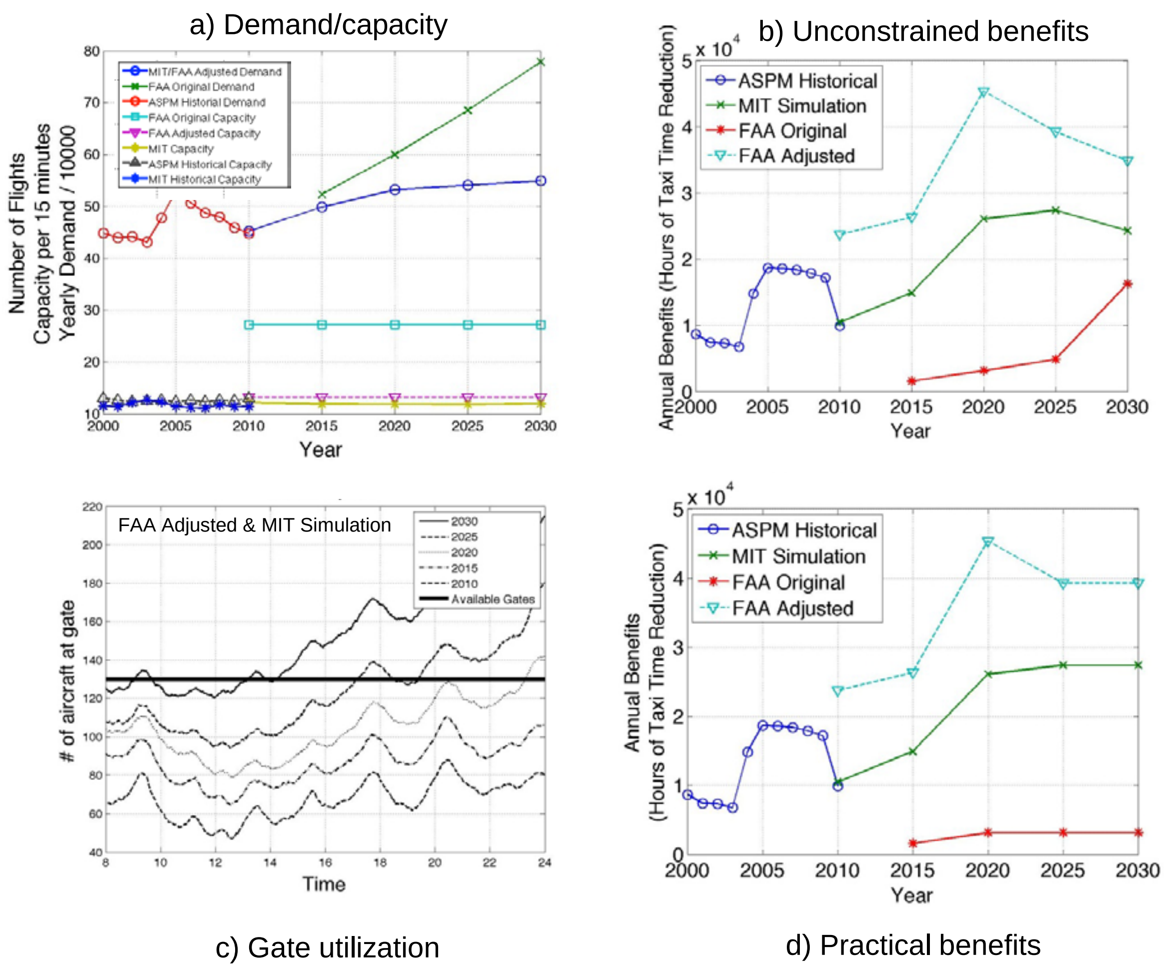


Figure 20: PHL Departure Metering Benefits Results

As shown in Figure 20, Philadelphia airport shows a medium level of benefits comparable to other airports in the study. The forecast demand had to be greatly trimmed due to a very high estimate of capacity. It is interesting to note that the FAA Adjusted level of benefits is above the MIT Simulation. This is due to anomalous behavior in the FAA taxi time simulation (see previous discussion). The FAA Original benefits behave as expected, so the additional trimming could have had unexpected effects on the FAA simulation. This could also be the reason behind 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. However, if the curve is taken as is, 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.

3.4.9 Aggregate Departure Metering Taxi Time Benefits

The aggregate departure metering benefits with respect to taxi time for each input are shown in Figure 21 for the unconstrained cases and Figure 22 for the practical cases. The airports with major contributions to the unconstrained benefits are ORD, ATL, and JFK in the MIT Simulation results, with the other five airports at a lower level. The largest contributor to the FAA Original results is JFK, while JFK, LGA, IAD and ATL are the biggest contributors to benefits in FAA Adjusted results. When gate constraints are accounted for, the practical benefits are significantly lower at the constrained airports, as expected. The relative importance of the airports in each input case is largely consistent with the unconstrained results because even the constrained airports exhibit significant benefit values and the lower benefit airports are typically not gate constrained. However, there is some swapping in the relative benefits between airports, for example in the unconstrained case the ranked order of airports is ORD, ATL, JFK for the MIT Simulation results in 2030, but this changes to ATL, JFK, ORD once gate constraints are included. Note that for the FAA Adjusted case LGA becomes the dominant airport overall.

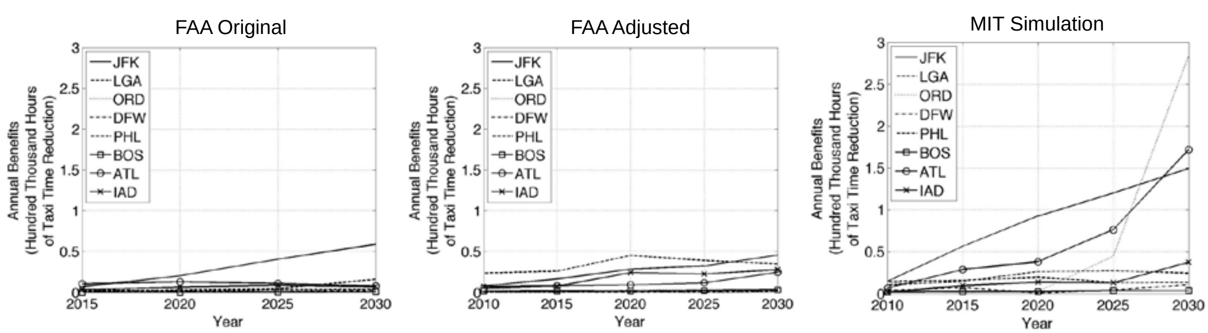


Figure 21: Aggregate Unconstrained Taxi Time Benefits

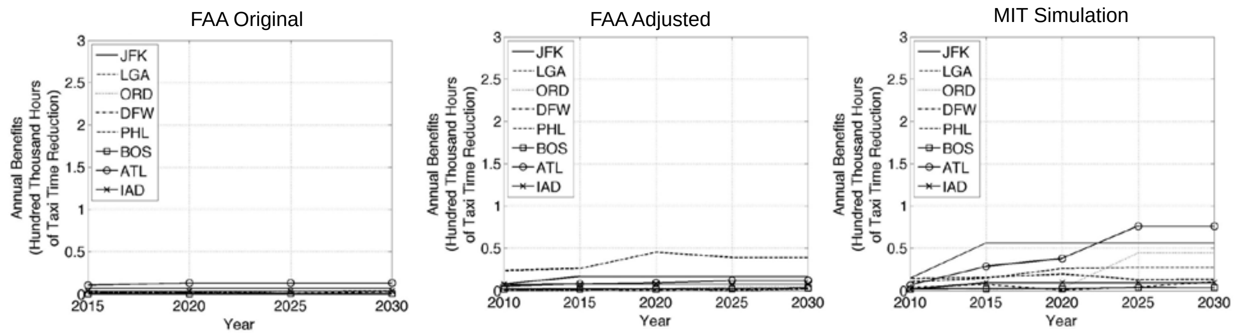


Figure 22: Aggregate Practical Taxi Time Benefits

3.4.10 Aggregate Departure Metering Fuel Burn Benefits

The aggregate departure metering benefits with respect to fuel burn for each input are shown in Figure 23 for the unconstrained cases and Figure 24 for the practical cases. The active taxi time benefits were converted to fuel burn savings by multiplying by ICAO-standard ground idle fuel flow rates for different aircraft types [ICAO (2011)] consistent with the fleet mix at different study airports. The specific fuel flow rates for the study airports are given in Table 2.

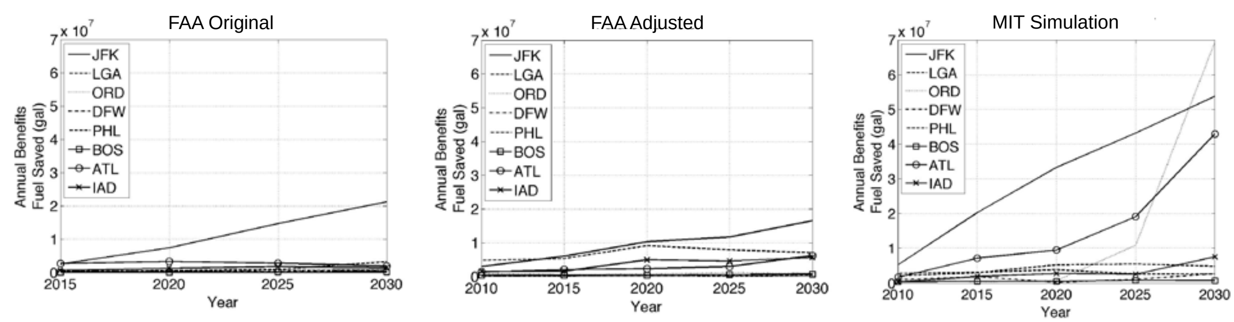


Figure 23: Aggregate Unconstrained Fuel Burn Benefits

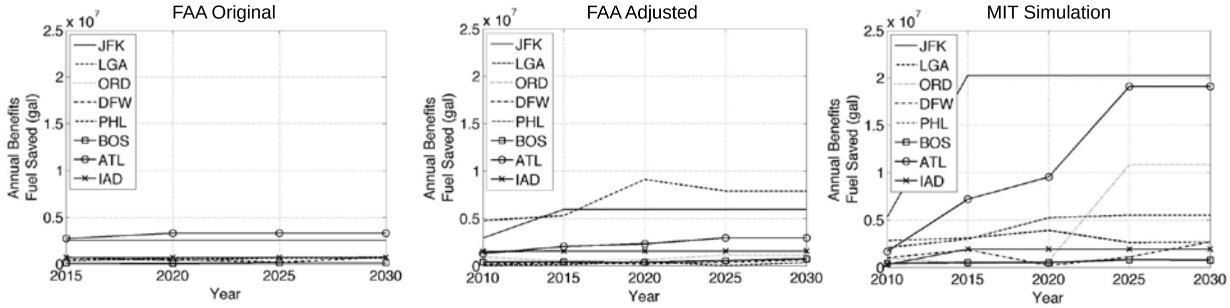


Figure 24: Aggregate Practical Fuel Burn Benefits

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

The relative contributions of the different airports to the fuel burn benefits are similar to those seen in the taxi time results. Observable differences are caused by fleet mix differences between airports. For example, the practical taxi time benefits for the MIT Simulation results in 2030 show most benefit from ATL and second most from JFK, while the order is switched in the fuel benefits because the fuel burn rate at JFK is much higher than ATL because of the larger fraction of wide-body aircraft at JFK.

3.5 DISCUSSION

3.5.1 Results Summary

Tables 3–5 present summaries of the taxi time and fuel benefits by airport summed over the 2015–2035 time period of interest to TFDM. The fuel burn savings were monetized by assuming a fuel price of $2.43/gallon as recommended by the FAA for investment analysis purposes [FAA/ATO (2011)]. Across the eight study airport, the unconstrained benefits estimates over the 2015–2035 time period range from 1.5 million hours taxi time and 448 million gallons of fuel (with a value of $1.1 billion) for the FAA

Original input case, to 8.5 million hours taxi time and 2.3 billion gallons of fuel (with a value of $5.6 billion) for the MIT Simulation input case. The practical benefits estimates over the 2015–2035 time period range from 727 thousand hours taxi time and 188 million gallons of fuel (with a value of $456 million) for the FAA Original input case, to 4.3 million hours taxi time and 1.1 billion gallons of fuel (with a value of $2.8 billion) for the MIT Simulation input case.

Table 3: Departure Metering Benefits Summary 2015–2035 (FAA Original Input Data)

Airport	Unconstrained Case			Practical (Gate Constrained) Case		
	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)
ATL	218	55	133	269	67	164
BOS	14	3	7	11	3	6
DFW	41	11	26	41	11	26
IAD	143	29	70	71	14	35
JFK	802	288	700	148	53	129
LGA	65	13	31	65	13	31
ORD	61	15	36	59	14	35
PHL	174	35	85	61	12	30
TOTALS	1,519	448	1,089	727	188	456

Table 4: Departure Metering Benefits Summary 2015–2035 (FAA Adjusted Input Data)

Airport	Unconstrained Case			Practical (Gate Constrained) Case		
	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)
ATL	328	82	199	224	56	136
BOS	54	12	29	54	12	29
DFW	35	9	22	35	9	22
IAD	483	97	236	164	33	80
JFK	723	260	632	348	125	304
LGA	28	6	14	28	6	14
ORD	70	17	41	80	19	47
PHL	781	157	382	816	164	399
TOTALS	2,501	639	1,554	1,748	424	1,030

Table 5: Departure Metering Benefits Summary 2015–2035 (MIT Simulation Input Data)

Airport	Unconstrained Case			Practical (Gate Constrained) Case		
	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)	Taxi Time (thousand hours)	Fuel Burn (million gallons)	Fuel Burn ($ millions)
ATL	2032	508	1236	1269	318	772
BOS	66	14	35	66	14	35
DFW	131	34	82	131	34	82
IAD	463	93	226	201	40	98
JFK	2429	873	2122	1185	426	1,036
LGA	320	63	154	320	63	154
ORD	2512	612	1488	597	146	354
PHL	506	102	248	531	107	260
TOTALS	8,458	2,300	5,590	4,301	1,148	2,790

The sections which follow address some of the uncertainties which remain in the analysis for some specific airports and these should be considered when interpreting the results.

3.5.2 Uncertainty in Runway Capacity/Performance

Primarily 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 (e.g., between achieved performance 2000 and 2010). 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 12. ORD, on the other hand, would likely not be able to sustain operations at the demand level of 2030 since the estimated 300,000 hours of gate hold translates to about 40 minutes for each flight. This is unlikely to be an acceptable level of gate hold and would likely translate to taxi times over 1 hour for all flights without metering.

3.5.3 Gate Utilization

Primarily 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 (Figure 25). Because the methodology behind the future schedules is somewhat unclear to MIT LL, 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 ATL does not follow the trend for the previous years, also because of a difference in distributions. The cause for the difference in this case is an imbalance between arrivals and departures of up to 10% in the years before 2030. Arrivals and departures are balanced in 2030.

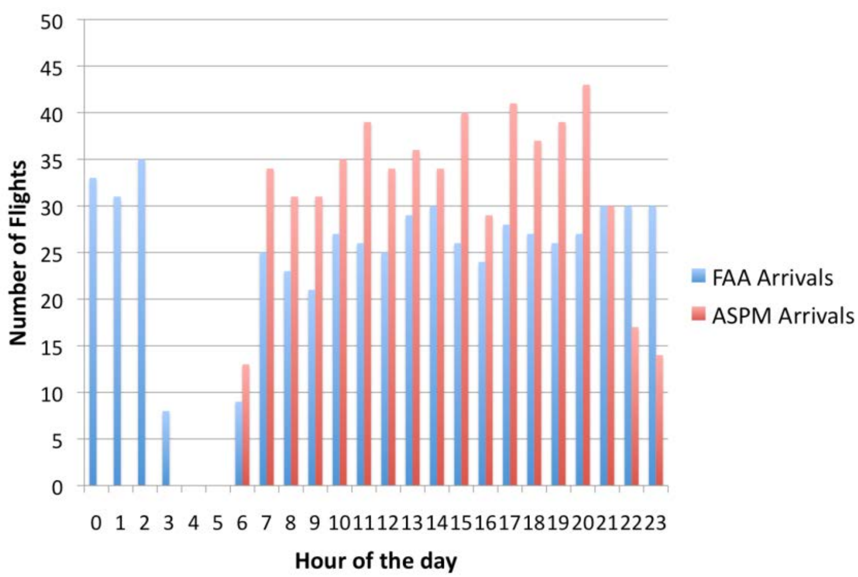


Figure 25: Comparison of Flights per Hour in 2010 at LGA

3.5.4 FAA Simulation Behaviors

Primarily Impacts: PHL

PHL displays unexpected behavior in the FAA Adjusted benefits by having higher benefits than the MIT Simulation, unlike any of the other study airports. It is believed that this is due to some unknown behavior causing the saturation curves to not match expectations. This is analogous to the LGA case that was shown in the results. In the left side of Figure 26, the blue (FAA) and red (MIT) curves are similar up until N = 30, where the blue throughput drops suddenly. Because the taxi time simulation for the FAA-based future year data is unknown, no hypotheses can be presented about why this occurs. For comparison, the ASPM saturation curve for 2010 is shown on the right side of Figure 26. This behavior further justifies the decision to use a separate taxi time simulation.

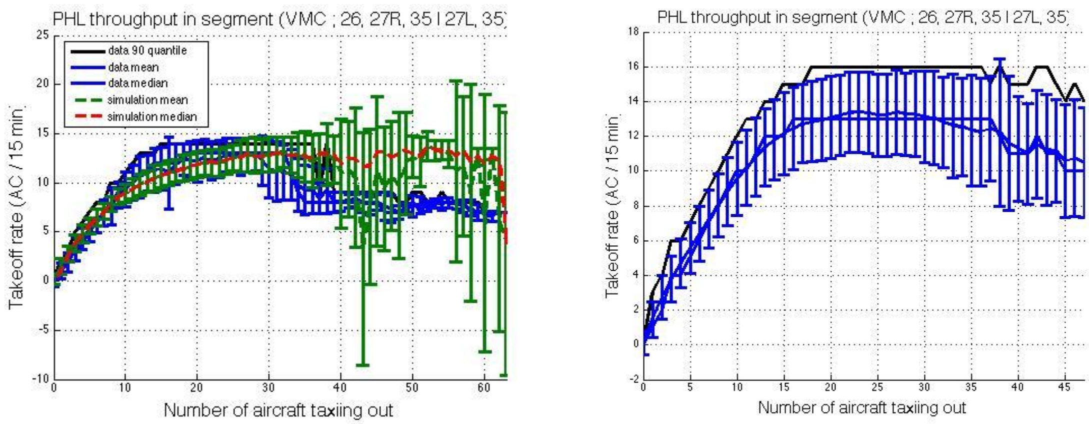


Figure 26: PHL 2010 Simulated (left) and ASPM (right). Throughput Saturation Curve Comparison

3.5.5 Extension of Findings to Other Airports

Potential benefits from departure metering are seen to vary greatly from airport to airport. This makes extending the benefits assessment to airports not covered in this study a challenge. Three possible methods are discussed here. One options is to weight the benefits of the eight study airports relative to the rest of the NAS. From ASPM, the eight study airports account for 41% of the taxi delay out of the total for the 43 TFDM airports in 2010. Hence there was 2.45 times more delay over all the 43 analysis airports compared to the eight study airports. Assuming this scaling factor remains constant over time (which is highly uncertain and would need further analysis to validate or refine), the estimated practical benefits using the MIT Simulation estimate would be $6.8 billion over the 43 airports. Other approaches could include using engineering judgement to place the study airports into clusters based on appropriate characteristics and then map a wider set of airports to those clusters. Benefits levels could then be cluster-specific. For example, a high benefits cluster could correspond to ATL, JFK, and ORD levels of benefits, the medium tier could correspond to PHL, LGA, DFW, and IAD and the low tier correspond to BOS. A third method is more involved and would identify and calculate a metric (most likely % of capacity used) to make a regression of the benefits. Other airports would then fall on that line. It is currently unknown how well behaved this regression would be.

Further work is required to assess which approach is most appropriate for extrapolation of benefits beyond the study airports and this would be high value for future work (see Section 7).

This page intentionally left blank.

4. SEQUENCE OPTIMIZATION

4.1 BACKGROUND

The goal of aircraft sequencing is to provide controllers a sequence advisory for both departures and arrivals needing to use a given runway. The problem formulation adopted for the TFDM benefits assessment application for optimizing aircraft sequences is largely based on the approach developed by NASA Ames: the Spot and Runway Departure Advisor (SARDA) [Gupta et al. (2009, 2010)]. The approach aims to control the gate release times of individual aircraft with a goal of improving throughput, reducing delays, and reducing fuel burn. Other surface optimization approaches have been proposed in the literature [Lee et al. (2010), Solveling et al. (2010)]. The advantage of the SARDA algorithm is that it controls only pushback times for departures and runway crossing times for arrivals while other algorithms attempt to schedule aircraft at a large number of control points on the surface. In addition, SARDA has undergone initial testing in both computer and human-in-the-loop simulations. Consequently, SARDA is considered to be more feasible for near-term operational applications than the other, more complex approaches.

We consider two types of sequence optimization. The first one is a basic version that provides sequencing advisories to the Local Controller for aircraft already at or near the departure runway. Because aircraft are to be sequenced at the runway, it was assumed that at least two sub-queues were available for departure re-sequencing. The second type of sequence optimization is a more advanced version that provides an advisory to the Ground Controller for gate[1] release times of individual departures. The basic form is complementary to departure metering discussed in the previous section such that their benefits can be considered additive. The more advanced form can be considered a substitute for departure metering in the longer term given that it is combining gate hold and sequencing functions. At the core of both methods is of an optimization program that minimizes a given objective function by scheduling the runway entry times for departures and arrivals while taking into account the separation needed between individual flights and the required time windows for individual departures. Both algorithms assume that the controllers will adhere to the minimum separation requirements between the aircraft and to the minimum observed times for the separations between departures and arrivals[2]. The second algorithm also assumes that the departures can take their delay at the gate resulting in reduced fuel burn and unimpeded

---

[1] At some airports (e.g., DFW) gate pushbacks are controlled by the airline ramp towers instead of the air traffic controllers. In such cases air traffic controllers schedule spot leave times instead of pushback times following advanced sequence optimization advisories.

[2] A range of minimum separation requirements has been tested as part of parametric analysis for basic optimization.

taxi time travel to the runway. The algorithm also assumes that the departures are able to maintain the appropriate speed so that their taxi times are equal to the times used by the algorithm (e.g., typical unimpeded times). In such a case, the gate release sequence would produce the desired runway entry sequence and no secondary optimization for the Local Controller would be required. Because it is simpler to implement in real-time, the basic version of sequence optimization is more applicable for near-term benefits assessment. Sequence optimization results in delay savings and reduced fuel burn. The delay savings are achieved by tightening separation constraints between the individual aircraft and sequencing the departures and arrivals. In the advanced version of sequence optimization additional delay savings are achieved since the departures travel unimpeded to the runway. The fuel savings are achieved through the reduction of taxi times and by holding the departures at the gate with engines off in the advanced sequencing formulation.

4.2 ANALYSIS METHODOLOGY

The methodology for the benefits assessment of both sequence optimization capabilities is illustrated in the figures below.

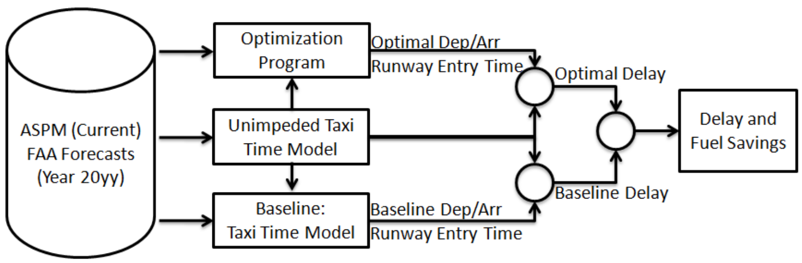


Figure 27: Advanced Sequence Optimization Analysis Methodology

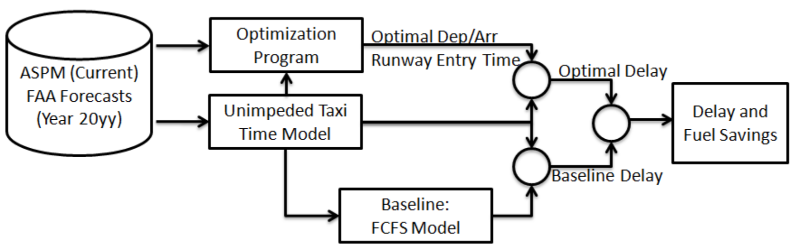


Figure 28: Basic Sequence Optimization Analysis Methodology

At the core of the analyses is an optimization that schedules runway entry times for each departure and arrival. The delay savings are then computed by comparing the optimized times to those produced by the baseline sequence. The baseline sequence for the advanced version required development of a taxi time model, whereas the basic version assumes modified first come first serve operations. The delay is defined as the time between the optimal or baseline runway entry time for each flight and an unimpeded time. The unimpeded time is the earliest time a particular aircraft can enter a runway. It is defined as the 10th percentile of the typical time that it takes an aircraft to travel from a particular terminal to a runway when there are no other aircraft on the surface. These times were derived from historical analysis of ASDE-X surveillance data. For future years, we assume that the surface unimpeded times will remain the same.

The following measurable benefit metrics are evaluated between the optimized and baseline behaviors:

1. The potential delay savings.

2. Fuel burn savings due to delay savings. For the advanced sequencing version, additional fuel burn savings are achieved since departure delay is absorbed at the gate instead of waiting in the active movement area. Note that for the case study of DFW these additional gate-hold fuel burn savings were not included because the aircraft are typically held at the spot with engines on.

As was described in the section above, two types of sequence optimization algorithms are evaluated. The baseline for the advanced sequence optimization version requires development of the taxi time model that would describe as-is operations. We compute the benefits of advanced sequence optimization for only one runway configuration: a departure runway with arrival crossings as exemplified by Dallas/Fort Worth (DFW) runway 17R operations.

4.2.1 Optimization Program

The optimization program used in this work is a mixed integer linear program [Gupta et al. (2009, 2010)]. It is used to schedule start-roll times for departures and start-cross times for arrivals which in turn define the sequence of aircraft using the runway for the first two runway configurations. The goal is to satisfy operational constraints while optimizing a certain objective function.

Several objective functions were compared to evaluate the trade-offs in achieved benefits, including minimizing average delay, minimizing maximum delay and minimizing the time the last aircraft entered the runway (serves as proxy for maximizing throughput). It was found that minimizing a weighted combination of average departure and arrival delays produced the greatest benefits in delay savings for the first runway configuration. This objective function also resulted in good throughput and maximum delay performance whereas the other two objective functions produced higher average delays. The weights used in the results below had the departure delay weighted ten times as much as the arrival delay. The results were sensitive to the weight ratio only if the departure delay weighting was increased substantially. In the limit of zero weight on arrival delay, the arrival delay savings were zero. In comparison, if the departure delay weighting was reduced, the arrival delay savings changed little. Again, delay is defined as the difference between the time an optimizer schedules runway entry for a particular aircraft and the earliest time that a particular aircraft can enter a runway.

The following constraints were used by the optimizer:

1. The earliest time a particular aircraft can enter a runway. For departures, the earliest time was set equal to pushback (or spot leave) time plus an unimpeded taxi time. For arrivals it was set equal to wheels-on time or inter-runway-entry time depending on the scenario. Since spot enter times were not available for future years, scheduled gate out times were used in the analysis of departures. Scheduled gate in times were used for the analysis of arrivals since inter-runway entry times were not available in the forecasts.

2. Separation requirements between departures on a single runway. Minimum separation requirements between individual departures are captured in the table below [Gupta et al. (2009, 2010)]. Minimum separation time for two consecutive departures going through the same departure fix was 90 seconds (derived from 5 nm minimum separation requirement).

Table 6: Runway Separation Requirements for Departures [sec]

		Trailing			
		Small	**Large**	**Heavy**	**B757**
Leading	**Small**	60	60	60	60
	Large	90	60	60	60
	Heavy	120	120	90	120
	B757	120	90	90	90
Departure fix minimum is 90 secs					

3. Separation requirements between departures and arrivals for the DFW 17R runway configuration with arrival crossings. The following additional minimum observed separations times were used (based on 2009/2010 data analysis at DFW). These were assumed not to change in the future.

    a. Minimum observed separation time between arrivals in the same arrival crossing queue: 10 seconds.

    b. Minimum observed separation time between arrivals in different arrival crossing queues: 2 seconds. This time accounts for the time it takes a controller to issue a command to two different crossing aircraft.

    c. Minimum observed separation time between a departure following an arrival: 12 seconds = 10 seconds for an arrival to cross runway 17R plus 2 seconds for a controller to issue a departure clearance.

d. Minimum observed separation time for an arrival following a departure: [30 35 40 45 52 60] seconds depending on which runway crossing an arrival is using. For example, the typical separation time between an arrival and the following departure at the top of the runway was 30 seconds while the separation time further down the runway can increase to 60 seconds. While these times vary with the aircraft size, this variability was not included in the current formulation due to the lack of the appropriate data analysis.

4. Separation requirements between arrivals on a separate runway for the configuration with two intersecting runways. The separation requirements for arrivals are shown in Table 7 below [Balakrishnan & Chandran (2010)]. FAA ASPM wheels-on and wheels-off data was used to find the typical waiting time between departures and arrivals on separate runways during congested periods. It was assumed that a departure flight could start its take-off roll 25 seconds after an arrival's landing and an arrival could land 55 seconds after a departure began its take-off roll.

Table 7: Runway Separation Requirements for Arrivals [secs]

		Trailing			
		Small	**Large**	**Heavy**	**B757**
Leading	**Small**	82	69	60	60
	Large	131	69	60	60
	Heavy	196	157	96	96
	B757	196	157	96	96

5. Flight-specific constraints such as target times of departures imposed from traffic management initiatives, e.g., expected departure clearance times (EDCTs) or calls for release (CFRs). These account for downstream constraints, for example, in the overhead stream, congested en route sectors or at destination airports. While sequence optimization algorithm can easily accommodate these constraints, they were not used due to the lack of input data.

The algorithm's planning horizon was restricted to fifteen minutes because forecast accuracy decreases further into the future and to keep computational times reasonable. At some parts of the day, the demand may exceed the available capacity and not all of the flights can be accommodated during the allocated fifteen minute time period. In such cases for the two scenarios without intersecting runways, some of the runway entry times suggested by the optimizer were carried over to the next time period and the optimizer preserved those times while scheduling the flights from the next fifteen minute bin afterwards. For the intersecting runways scenario, there was an additional constraint on arrival landing times. It was assumed that tower controllers have little control over the landing times and sequences of arrivals except for slowing them down by at most 30 seconds while they were airborne without altering the landing order. This number was suggested by the controllers we consulted.

It is important to note that this optimization algorithm assumes that all of the forecasts are perfect. However, in real operations there are uncertainties associated with forecasting pushback times, arrival times and taxi times. These uncertainties are not explicitly considered by this deterministic algorithm. Consequently, the results below provide an upper bound on potential sequencing benefits. Also, it is important to note that the computational time of this algorithm can be problematic for real-time implementation at higher demand levels. However, various successful techniques can be utilized to speed it up. They include formulating the optimization as a dynamic programming problem or using a combination of mixed integer and heuristic approaches. These techniques were not explored as part of this benefits analysis but could be appropriate options to consider for future analyses.

4.2.2 Baseline Model

As can be seen in the methodology figures, the benefits were computed by comparing the optimized runway entry times for each aircraft to the baseline. First we describe the taxi time model used as a baseline in advanced sequence optimization.

The baseline runway entry sequence is obtained from a taxi time model that was derived using ASDE-X flight track data over a set of thirteen VMC days in 2009/2010. The model was designed to predict the average time an aircraft would spend taxiing under current operating conditions given its originating terminal, runway, aircraft type and congestion on the surface. It was assumed that the taxi time model remains applicable even if the demand level is increased. At DFW, the demand grows by a factor of 1.5 in 2030 compared to 2010 traffic levels.

The departure taxi time model consists of two components: taxi time to the departure queue plus taxi time through the queue. The following variables were used to build the runway 17R taxi time model for departures: spot group of individual departures, aircraft weight class, number of aircraft waiting in 17R departure queue and number of arrivals that landed on runways 17C and 17L. Departure and arrival congestion variables were also used by the arrival wait time model. The performance of the models is summarized in Table 8 below, tested on three VMC days at DFW.

Table 8: DFW 17R Baseline Model Prediction Errors Tested on Three VMC Days

	|Error| < 1.0 min	|Error| < 1.5 mins	|Error| < 2.0 mins	Median taxi time
Spot to Queue Taxi Time Model	87%	96%	99%	3 mins
Queue Taxi Time Model	42%	64%	79%	7 mins
Arrival Time Model	86%	96%	99%	1 min

In comparison, the benefits for the basic sequence optimization does not require derivation of the detailed taxi time model as it assumes that the local controller can only sequence the aircraft queued up at the runway. The baseline model for basic optimization consists of a modified First Come First Served (FCFS) sequence that is separated using the same minimum separation requirements as the optimization program. In effect, basic sequence optimization evaluates the benefits of re-sequencing given the minimum separation constraints. For the runway configuration consisting of departures only, it was assumed that departures took off in the FCFS order which was consistent with current operations at DFW. The FCFS order was also assumed for departures in the intersecting runways configuration while the optimized and baseline arrival sequences were identical due to the constraints explained before.

For the runway configuration of a departure runway with arrival crossings, we assume that departures take off in the FCFS order. The arrivals cross in FCFS order as well with respect to other arrivals. For the interaction between arrivals and departures, we assume that controllers would wait to cross multiple arrivals simultaneously if arrivals were less than 20 seconds apart and in different crossing locations. In addition, controllers would cross arrivals that wait more than 40 seconds since arrival crossings have little impact on departure delay due to short runway crossing times.

4.2.3 Input Parameters

FAA ASPM database was used to generate inputs for current day operations and FAA forecasts were used for future years (the same future year data described in the Departure Metering section). For demand profiles we used LGA schedules for the case of intersecting runways and DFW demand profiles for the other two runway configurations. The schedules contained scheduled gate out for departures, scheduled gate in for arrivals and aircraft type information. Runway assignment, departure fixes and arrival crossing locations were not available directly from the schedules and had to be randomly generated. We assumed that current-day distributions associated with those variables remain the same in the future.

The arrival times for the case with intersecting runways required an adjustment. The published schedules had many flights arrive at regular time intervals such as on the hour. Since it is impossible for multiple flights to land at exactly the same time, the landing times were artificially separated so that the minimum separation requirements between arrivals were met.

4.3 AIRPORT SCOPE

Three runway configurations are considered for benefits assessment of basic sequence optimization:

1. A configuration with a departure runway with arrival crossings, as exemplified by DFW departure runway 17R and crossing arrivals from 17C and 17L. In this case multiple crossing points were available for arrivals enabling simultaneous crossings.

2. A configuration with a departure runway and no arrival crossings as exemplified by DFW runway 17R with arrivals using the perimeter taxiway.

3. A configuration with two runways that cross as exemplified by LGA runways 13 and 22. For this runway configuration we assumed departures and arrivals used separate runways.

4.4 RESULTS

4.4.1 Advanced Sequence Optimization at DFW 17R Departure with Arrival Crossings from 17C and 17L

Table 9 presents the results for the advanced sequence optimization case. The average daily delay savings are due to the reduction in taxi times because departures can travel to the runway unimpeded if they are released at optimal times from the gate (or spot for DFW operations), due to tightening up of separation constraints to the minimum standards and due to re-sequencing of departures and arrivals. The results show the average delay savings due to advanced sequence optimization over a period of twelve representative days at DFW. The 2010 demand levels used in the simulations consisted of 465 departures and 330 arrivals. These east-side operations respectively account for 55% and 35% of departure and arrival operations at DFW. According to the forecasts, the demand at DFW grows as follows 2010:2015:2020:2025:2030 = 1.0x:1.1x:1.2x:1.3x:1.5x. We assume that capacity remains the same over this time period. Note the decrease in departure delay savings in 2030 is due to a large rise in optimized delay (operating near capacity limits) compared to the delays obtained using the baseline taxi time model (see Figure 29 below). In other words, the delays grew from year to year in both optimized and baseline versions but the baseline rate of growth was lower at higher demand levels. The lower rate of growth of baseline delay can be attributed to the fact that the model was derived using data when demand was much lower than capacity and hence does not fully capture the nonlinear aspect of queuing when demand exceeds the capacity.

Table 9: Average Daily Savings due to Advanced Sequence Optimization for DFW 17R Operations with Arrival Crossings from 17C and 17L

Year	Departures		Arrivals	
	Av. Daily Delay Savings (hours)	Av. Daily Fuel Burn Savings (gallons)	Av. Daily Delay Savings (hours)	Av. Daily Fuel Burn Savings (gallons)
2010	16.8	4,319	3.4	874
2015	19.8	5,091	3.8	977
2020	23.0	5,914	4.6	1,183
2025	24.5	6,299	5.9	1,517
2030	18.2	4,679	9.9	2,545

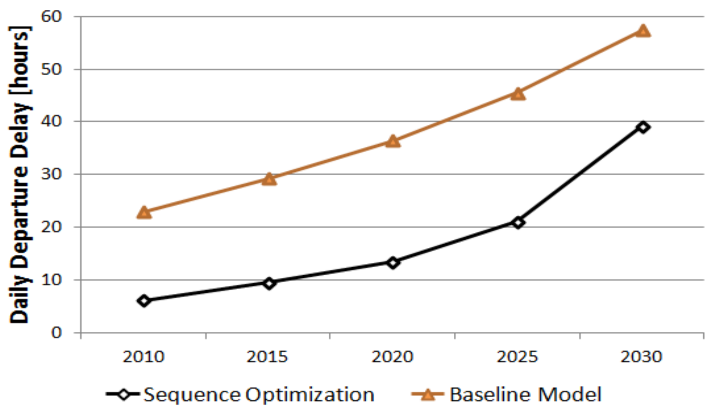


Figure 29: Comparison of Daily Departure Delays for Advanced Sequence Optimization

Fuel burn savings due to reduction in delay plus departures incurring delay at the gate were computed using these formulae:

$$Fuel\ Savings\ (Departures) = Delay_{Baseline} * FuelBurn_{Taxi} - Delay_{Optimal} * FuelBurn_{Gate}$$
$$Fuel\ Savings\ (Arrivals) = (Delay_{Baseline} - Delay_{Optimal}) * FuelBurn_{Taxi}$$

Since at DFW air traffic controllers cannot control the gate pushback times, the departure and arrival fuel savings were computed using:

$$Fuel\ Savings = (Delay_{Baseline} - Delay_{Optimal}) * FuelBurn_{Taxi}$$

At DFW, the airport-specific average fleet fuel burns are .2214 kg/s which converts to 257.1 gal/hr using the 3.1 kg/gallon density as was described in the departure metering section.

4.4.2 Basic Sequence Optimization at DFW 17R Departure with Arrival Crossings from 17C and 17L

Table 10 presents the results for the basic sequence optimization case. Since there is no taxiing, the delay savings of basic sequence optimization are due to re-sequencing of aircraft given the minimum separation requirements used both in the baseline FCFS sequence and the optimization program. Fuel savings were computed by multiplying the total delay savings by the DFW fuel burn rate of .2214 kg/s as before.

Parametric analysis was also performed to check the sensitivity of the results to changes in the input parameters. Analysis showed that the results were not sensitive to changes in separation times between arrivals and departures. However, the changes in delay savings were substantial when the separation requirements between departures were changed by plus or minus 20% (i.e., a factor of 0.8 or 1.2 on the baseline values given in Table 6), see Table 11. As opposed to the analyses of advanced and basic sequence optimization delays savings presented above, parametric analysis was done using demand profiles for only one sample day (March 9th) and multiple simulation runs to obtain statistically significant results.

Table 10: Average Daily Savings due to Basic Sequence Optimization for DFW 17R Operations with Arrival Crossings from 17C and 17L

Year	Departures		Arrivals	
	Av. Daily Delay Savings (hours)	Av. Daily Fuel Burn Savings (gallons)	Av. Daily Delay Savings (hours)	Av. Daily Fuel Burn Savings (gallons)
2010	0.3	77	1.1	283
2015	0.9	231	2.9	746
2020	1.9	489	4.1	1,054
2025	3.8	977	5.9	1,517
2030	11.7	3,008	9.9	2,545

Table 11: Variability in Departure and Arrival Delay Savings due to Basic Sequence Optimization for DFW 17R with Arrival Crossings as a Result of Changing the Minimum Separation Requirements (Results for One Sample Day)

Year	Av. Daily Departure Savings (hours)			Av. Daily Arrival Savings (hours)		
	0.8x Baseline Separations	1.0x Baseline Separations	1.2x Baseline Separations	0.8x Baseline Separations	1.0x Baseline Separations	1.2x Baseline Separations
2010	0.2	0.3	0.5	0.5	1.1	2.1
2015	0.6	0.8	1.4	1.6	2.9	5.4
2020	1.6	2.5	4.4	2.1	4.4	9.4
2025	3.2	4.3	8.2	2.6	6.1	15.2
2030	6.8	12.9	25.2	3.5	10.1	30.0

In addition, the sensitivity of the results to changes in departure fix use was explored. In the results presented below, eight departure fixes were used for 17R departures with 90 second minimum separation between two consecutive aircraft going over the same fix. It is referred to as Proc. 1 in Table 12. The results also illustrate an alternate departure procedure (Proc. 2) that is also used DFW. Instead of eight separate departure fixes, two departure headings are used and if two consecutive aircraft are on divergent headings then the runway separation requirements are reduced.

Table 12: Variability in Departure and Arrival Delay Savings due to Basic Sequence Optimization for DFW 17R with Arrival Crossings as a Result of Changing the Departure Fix Procedures (Results for One Sample Day)

Year	Av. Daily Departure Savings (hours)		Av. Daily Arrival Savings (hours)	
	Proc. 1	Proc. 2	Proc. 1	Proc. 2
2010	0.3	-0.3	1.1	0.6
2015	0.8	-0.2	2.9	2.0
2020	2.5	0.2	4.4	2.9
2025	4.3	0.8	6.1	3.8
2030	12.9	3.5	10.1	5.5

In this analysis we assumed a 25% reduction of the minimum separation requirements presented in Table 6. Given that there were only two departure headings, this procedure effectively reduced the average separation requirements between departures resulting in the reduction of delay savings due to basic sequence optimization. In the limit, sequence optimization would provide no benefit if the separation requirements between departures were zero. Note that while there are negative delay savings for departures in 2010 and 2015 using the alternate departure procedures, there is an overall benefit of using basic sequence optimization when departure and arrival savings are combined.

More than 90% of aircraft at DFW in current operations are classified as "Large" weight category which limits re-sequencing opportunities for the optimizer given the separation requirements vary by weight class. Parametric analysis was performed to explore re-sequencing benefits by artificially varying the aircraft fleet mix so that the results of the benefits analysis could be extrapolated to other airports. In addition to the fleet mix, the separation requirements matrix for departures was also varied by scaling it uniformly while all other parameters were held constant. The complete results for each simulation year are presented below and sample results for year 2015 in graphical form are presented in Figure 30. The fleet mixes considered in the simulation are as follows and represent the current-day fleet mixes at DFW (Fleet Mix I), SFO (Fleet Mix II) and JFK (Fleet Mix III), see Table 13.

Table 13: Description of Fleet Mixes Used in Parametric Analysis

Fleet Mix	Small	Large	Heavy	B757
I	1%	90%	4%	5%
II	15%	60%	15%	10%
III	2%	60%	29%	9%
IV	25%	25%	25%	25%

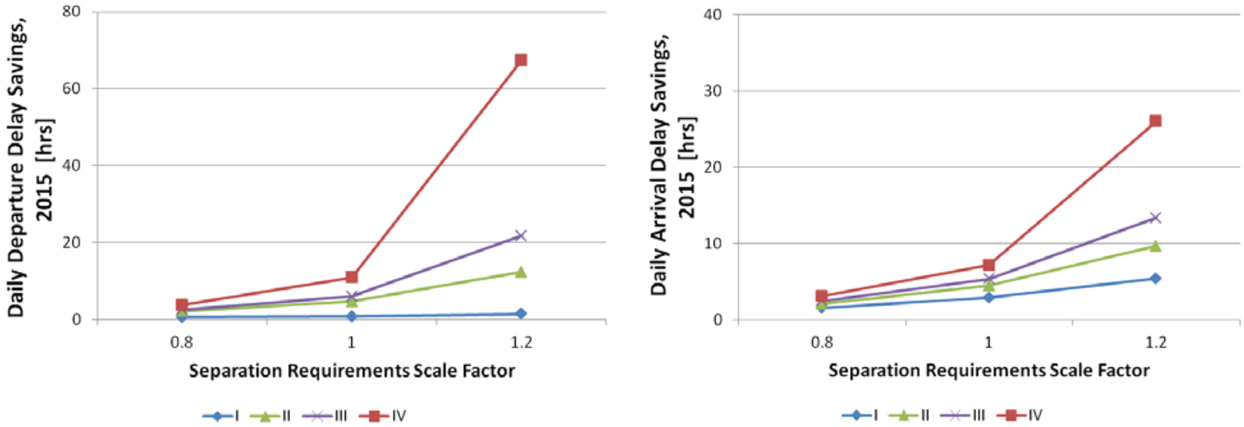


Figure 30: Variability in Delay Savings due to Basic Sequence Optimization Using 2015 Demand for DFW 17R with Arrival Crossings as a Result of Changing the Fleet Mix and Separation Requirements between Departures

As can be seen in the results in Table 14, the benefits of sequence optimization vary substantially when the fleet mix is changed.

Table 14: Summary of Delay Savings due to Basic Sequence Optimization for DFW 17R with Arrival Crossings as a Result of Changing the Fleet Mix and Separation Requirements

		Departures			Arrivals		
		Separation Requirement Factor			Separation Requirement Factor		
	Fleet Mix	0.8	1	1.2	0.8	1	1.2
2010	I	0.2	0.3	0.5	0.5	1.1	2.1
	II	1.0	1.9	4.7	0.9	2.0	4.5
	III	1.0	1.9	5.8	1.1	2.4	5.3
	IV	1.6	3.8	14.2	1.5	3.4	7.7
2015	I	0.6	0.8	1.4	1.6	2.9	5.4
	II	2.2	4.6	12.2	2.1	4.5	9.7
	III	2.5	5.9	21.7	2.5	5.3	13.3
	IV	3.8	10.9	67.4	3.1	7.2	26.0
2020	I	1.6	2.5	4.4	2.1	4.4	9.4
	II	5.7	14.9	71.4	3.3	7.6	26.7
	III	6.6	19.9	119.3	4.0	9.7	42.5
	IV	8.9	40.6	N/A	5.2	14.9	N/A
2025	I	3.2	4.3	8.2	2.6	6.1	15.2
	II	9.6	27.7	144.4	4.5	11.7	45.2
	III	12.3	41.1	188.5	5.6	16.0	68.7
	IV	16.1	98.5	N/A	7.2	26.5	N/A
2030	I	6.8	12.9	25.2	3.5	10.1	30.0
	II	22.5	103.5	N/A	7.1	25.2	N/A
	III	28.7	171.5	N/A	8.9	39.5	N/A
	IV	47.7	N/A	N/A	13.7	N/A	N/A

The opportunities for re-sequencing and hence the benefits are largest when all types of aircraft are represented equally (Fleet Mix IV). The benefit of re-sequencing also increases when the inter-departure separations are increased. Note that in Table 14 some of the values are missing. The reason is that at higher demand levels (e.g., in 2020) demand often exceeds theoretical throughput resulting in operationally unrealistic median baseline delays of greater than 40 minutes, so these results have been omitted.

Table 15 shows the total delays and fuel burn savings due to sequence optimization at DFW east side operations (departure runway 17R, arrivals on 17C and 17L). Currently, east side operations account for about 55% of all departures and 35% of all arrivals at the airport. The delay savings were calculated by linearly interpolating results for years 2015, 2020, 2025 and 2030 to obtain the potential daily delay savings at other years. Delay savings in years 2031–2035 were taken to be equal to 2030 values. The average daily delay savings computed over the twelve representative days in the FAA schedule data were then multiplied by 365 to get yearly benefits for advanced and basic optimizations with Fleet Mix I. Since the effect of changing fleet mixes was tested only on one sample day (03/09), the delays savings for that day were multiplied by 365 to get the total savings due to basic optimization for Fleet Mix II and III. Fuel burn savings were computed using airport-specific fuel burn rates for DFW, SFO, and JFK (.2214 kg/s, .2641 kg/s, .3096 kg/s) respectively for Fleet Mix I, II, and III. These fuel burn rates were multiplied by the delay savings to get the fuel burn savings since advanced sequence optimization at DFW does not include gate holds and basic optimization only re-sequences aircraft at the runway. DFW-wide results estimates can be estimated by doubling the benefits estimates given in Table 15 to account for operations on both sides of the airport.

Table 15: Total Delay and Fuel Burn Savings for DFW East Side Operations for Years 2015–2035 due to Basic Sequence Optimization for a Single Departure Runway With Arrival Crossings

Optimization	Fleet Mix	Total Passenger Departure Delay Savings (thousand hours)	Total Passenger Arrival Delay Savings (thousand hours)	Fuel Savings (million gallons)	Fuel Savings ($millions)
Advanced	I	162	52	55	134
Basic	I	46	51	25	61
	II	385	114	153	372
	III	619	168	283	687

4.4.3 Basic Sequence Optimization at DFW 17R Departures Only

This scenario consists of a single departure runway with no arrival crossings. DFW departure demand on 17R was used for this analysis. The results are given in Table 16 below.

Table 16: Comparison of Departure Delay Savings due to Basic Sequence Optimization for Two Runway Configurations

Year	Av. Daily Departure Savings (hours)	
	Departures with Arrival Crossings	**Departures without Arrival Crossings**
2010	0.3	0.4
2015	0.9	0.9
2020	1.9	2.1
2025	3.8	3.7
2030	11.7	9.7
Total savings for DFW 17R departures-only operation 2015–2035: 40,000 hrs, 10 million gallons fuel		

The results show that, compared to the scenario of departures with arrival crossings, there are fewer opportunities for sequencing in this scenario resulting in decreased delay savings everything else being equal. The total delay and fuel burn savings for years 2015–2035 computed similarly to the way described in the section above are 40,000 hours of departure delay savings and 10 million gallons in fuel burn savings compared to the much higher values in Table 15.

4.4.4 Basic Sequence Optimization for Intersecting Runways

LGA demand schedules were used from 2010 to perform current day benefits analysis of basic sequence optimization for an intersecting runway case typified by the LGA configuration. Given that 2010 demand levels often exceed the stated theoretical capacity, many of the scheduled flights did not fit into the allocated time bins used for optimization. These surplus flights needed to be re-optimized, and as such the benefits of the optimization on the previous time bin could not be achieved. This resulted in FCFS sequences performing better on a given set of isolated flights. We also found that the results were very sensitive to variations in separation requirements between flights compared to the two other runway configurations described above. For these reasons, no results are presented for this runway configuration. Further analysis should be performed to provide realistic benefits of basic sequence optimization for the intersecting runways scenario, and this is another option which could be considered for future analysis.

4.5 DISCUSSION

We explored the benefits of two types of sequence optimization strategies. The advanced version aims to control gate/spot push-backs of aircraft to accomplish an optimal departure sequence, while the basic version provides advisories to improve the departure sequence of aircraft already at or near the runway. Both versions require accurate predictions of unimpeded taxi times, pushback times and wheels-on times for arrivals. For example, basic sequence optimization needs forecasts of aircraft runway ready times to perform optimization on a selected time interval. In addition, the advanced version requires that the departures are able to hold at the gate and then travel unimpeded to the runway. It is important to note that the types of accuracies required by deterministic optimization described above have not yet been achieved by the tools in the field. In order to make sequence optimization algorithms work in an operational setting, two research avenues should be pursued. First, better models for taxi times, push back times and wheels-on times should be developed. Second, stochastic optimization techniques need to be used so that the uncertainty associated with those times is explicitly modeled. Otherwise, given the uncertainty present in the input parameters, the majority of the estimated benefits due to deterministic sequence optimization might be eliminated once it is tested in the field. As was mentioned earlier, due to the lack of explicitly modeling the stochastic effects present in the actual operational environment, the delay savings presented above provide an upper bound on potential benefits. The results of basic sequence optimization for a single departure runway case with and without arrival crossings can be used to extrapolate the benefits for other similar configurations. Since parametric analysis was performed for various fleet mix scenarios, the results are applicable to other airports as long as the demand to capacity ratio at those airports is similar to the one at DFW.

This page intentionally left blank.

5. AIRPORT CONFIGURATION OPTIMIZATION

5.1 BACKGROUND

Airport configuration changes frequently occur at airports as a result of changing wind, weather and other operational conditions such as noise abatement requirements and configurations of neighboring airports. For example, there are federal, local, and operator limits to the tail wind, cross wind, and gust components that can be allowed for landing and departing aircraft. When these limits are exceeded, operations are moved to conforming runways. Boston Logan airport experienced nearly 1900 configuration changes in 2010, see Figure 31.

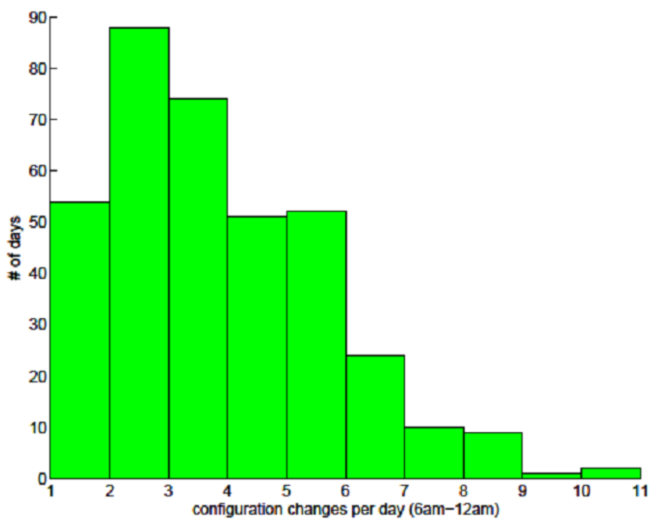


Figure 31: Histogram of Daily Airport Configuration Changes at Boston Logan Airport in 2010

Such configuration changes often result in a period of reduced capacity while ground and TRACON traffic are rerouted to align with the flows into and out of the new configuration. When these changes occur during periods of high demand, they can result in significant surface congestion with excess delay, fuel burn, noise and emissions. Consequently, the selection, timing and frequency of configuration transitions are important considerations in reducing delays and emissions around airports.

Several proposed methods seek to optimize the sequence of airport configurations. For the TFDM application, we extend the runway configuration optimization framework proposed by Bertsimas and

Frankovich [Bertsimas et al. (2011)], itself an extension of the airport capacity optimization framework introduced by Gilbo [Gilbo (1993, 1997)]. This approach was selected given that it builds upon the existing literature and accommodates operationally-relevant constraints (e.g., provisions for uneven prioritization between arrival and departures, variable time costs for transitions and considerations for neighboring airports). The approach and its predecessors rely on runway configuration a envelopes (RCCEs), outer convex envelopes to arrival vs. departure counts during regular time intervals. Each RCCE (illustrated in Figure 32) reflects the airport's ability to accommodate departure and arrival demand under various runway and wind directions, visual or instrument meteorological conditions (VMC or IMC), runway conditions and closures, noise abatement, surrounding airport configurations and other local conditions. Runway configuration sequences (and hence airport capacity) are therefore optimized over a given time horizon by minimizing un-served traffic count given demand levels and sequences of feasible runway configurations.

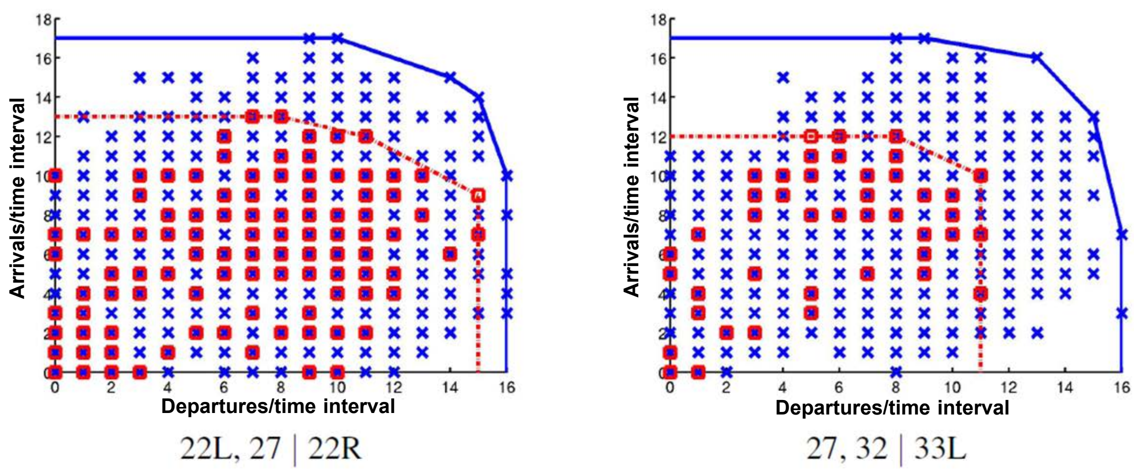


Figure 32: Sample BOS RCCEs. VFR and IFR Data Points Are Respectively Depicted in Blue and Red

5.2 ANALYSIS METHODOLOGY

The approach outlined in [Bertsimas et al. (2011)] has been adapted and extended to support the needs of the TFDM benefits analysis of the airport configuration optimization capability using the methodology illustrated in Figure 33. Currently, our RCCEs are generated from airport arrival and departure historic operational data as reported in the ASPM 2010 database in 15 minute time intervals for each operating configuration. Future year RCCEs (for 2015, 2020, 2025 and 2030) are estimated based on the FAA-sanctioned forecasts of future demand, capacity and schedules as described in the previous sections for the other IID analyses. To simplify the runway configuration definitions, the set of all possible configurations generated from past years' RCCEs, reported in ASPM for 2010 (68 for Logan) is reduced to a representative macro group (5 for Logan).

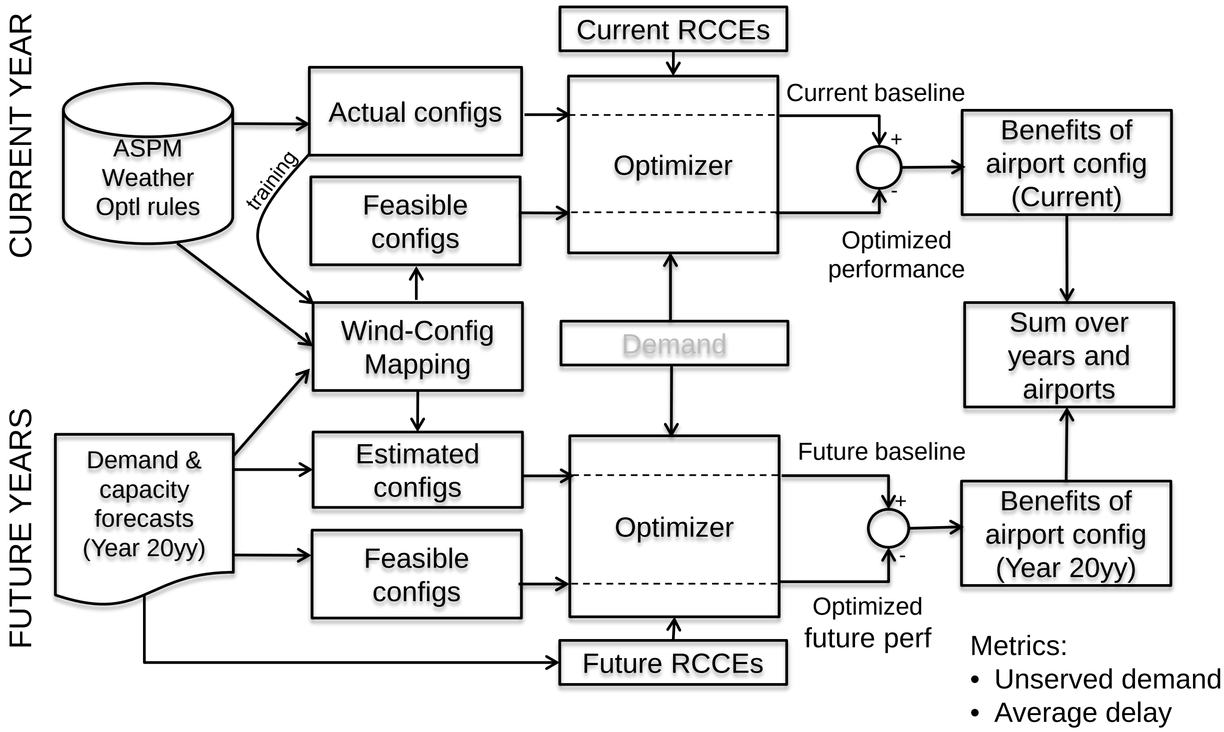


Figure 33: Airport Configuration Optimization Analysis Methodology

Historic data is parsed to identify attributable relationships between runway configuration, weather and other operational data. These relationships are subsequently leveraged to define the sets of feasible configurations. Each optimized configuration is selected from the set of feasible configurations at each time slot given weather and other input conditions. Input from air traffic control personnel at airports of interest provided additional operational validity to the models. This all-inclusive set of feasible configurations then provides the optimization engine with constraints. From these it considers how much of the predicted arrival and departure demands in each 15 minute intervals can be served by each configuration in the feasible set given the performance of each configuration and the loss in capacity associated with a configuration change. An optimization objective function of the form [Bertsimas et al. (2011)]:

$$\min \sum_t (c_t u_t + q_t v_t)$$

is applied where u_t = the number of arrivals which go un-served in time interval t; c_t = the cost of delaying arrivals during the time interval t; v_t = the number of departures which go un-served in time interval t; and q_t = the cost of delaying departures during the time interval t. This objective function represents the total (weighted) number of flights unable to complete their scheduled operations during the allocated time slots, i.e., the unserved demand. The weights are unit-less and represent the relative importance allotted to arrivals versus departures during the optimization. For example, c_t = 12 and q_t = 10 during time slot t signifies that the cost of moving arrivals out of their time slots is 1.2 times that of departures. Based on the relative weights and numbers of un-served arrivals and departures, an optimal sequence of RCCEs (and hence airport configurations) is suggested for the time interval considered. The performance of this optimized set of configurations in terms of throughput, delay, fuel burn, and emissions can then be compared to the baseline case in order to determine the benefits (if any) of the optimization capability.

The benefits of the runway configuration optimization tool are assessed through the following steps:

1. Establishing scheduled and actual arrivals and departures counts from the ASPM database or simulated forecasts

2. Computing a baseline performance of un-served traffic under given demand, weather and other operational conditions [D1]

3. Generating feasible airport configuration rules for mapping weather/operational conditions to configurations based on historical data

4. Generating optimal configuration sequences and computing associated operational performance [D2]

5. Assessing benefits = [D2] – [D1]

The process and its assumptions are described in the following text. Future years modeling assumes weather patterns similar to the year 2010 as well as the use of same physical runways in 2015, 2020, 2025, and 2030. Therefore, the feasible set of configurations at each given time slot is assumed unchanged in future years from the same 2010 date/time.

5.2.1 Data Sources

Archived meteorological data (METAR) from OGINET (both hourly and special advisories) as well as the FAA ASPM Performance Model (APM) database provided the critical data necessary for the study.

APM provided past statistics on actual arrival and departure demand as well as runway configurations usage per quarter hour. The FAA ASPM database yielded individual flight information for each airport during the years of interest. This data helped correlate artificially generated bin counts to individual flight delays. Simulated future year (select 2010–2030 days) individual flight data provided by the FAA provided statistics for estimating future benefits (the same data used in the other IID analysis areas previously described).

As expected, the simulated future traffic demand showed a monotonic increase from year-to-year. The traffic data sources are referenced as follows:

a. Individual flight statistics and counts from the FAA ASPM reporting 15 minute actual traffic and runway configuration usage
b. Simulated individual flights generated from the FAA for select past and future years

A comparison of the traffic from these two sources during 2010 in Boston is shown in Figure 34. We note the disparity between arrivals and departures in 2010 across data sources. Future demand increases monotonically as expected.

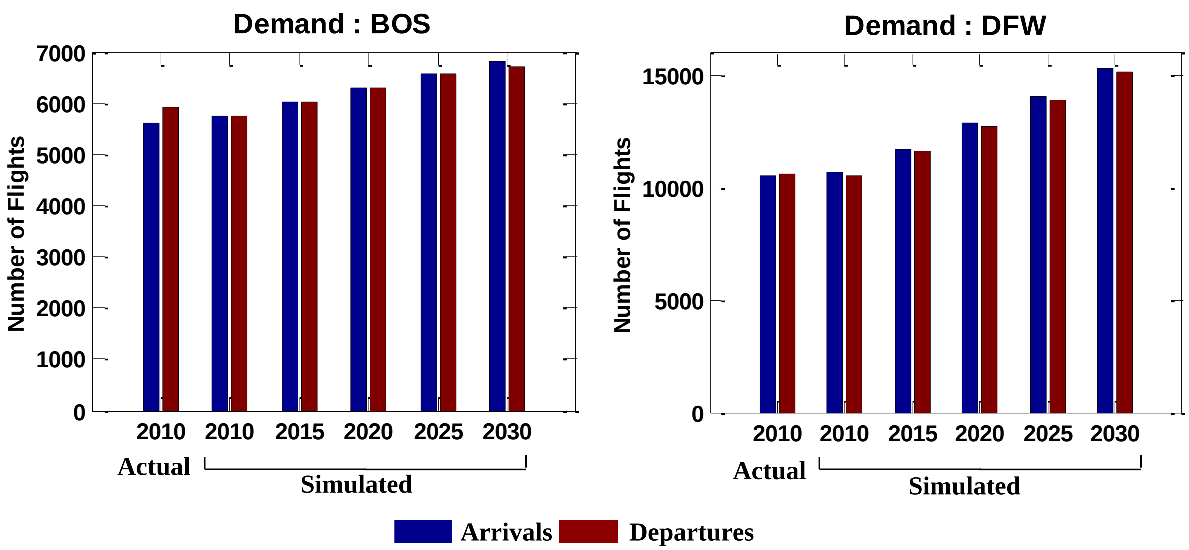


Figure 34: Actual and Forecast Traffic Comparison

In addition to published databases, this study included data from observations and interviews obtained during field visits to New York and Boston Centers, TRACONs and towers. These visits to the various towers in the North East provided important operational observations as well as controllers' perspective on the airport configuration management processes.

5.2.2 Baseline

Baseline delays solely attributable to airport configuration changes are difficult to decouple from the many other contributing factors which cause delay. Therefore, a relative measure of airport configuration benefits is adopted using the optimization algorithm to produce both the baseline and the fully optimized performance. A baseline reference of un-served operations is then established by restricting the optimization engine's configuration set at each time slot to the recorded configuration published in the ASPM database for that time slot. The optimization process computes a baseline count of un-served traffic due to this forced configuration sequence input. The relationship between airport configuration and weather (from METAR) and other operational conditions (from airport Standard Operating Procedures (SOPs)) is established from recorded ASPM data. The performance of the airport in terms of throughput, delay and fuel burn in this baseline case is assessed. A simple lookup table proved sufficient for approximating controller's behavior in the scope of this study.

5.2.3 Runway Configuration Capacity Envelopes

The methodology used in this study depends on the availability of runway capacity envelopes for the airports of interest, and the faithful capture of airport capacity performance statistics through these envelopes. These envelopes are therefore generated from historic data, with particular attention to years during which the airport operated at higher capacity. High capacity years provide a sense for the operational limits of the airports. In the case of Boston, the primary focus of this study, FAA ASPM data counts were used to generate the RCCE curves for VMC and IMC weather conditions. The higher traffic volume in 2001 (424,445) as opposed to 346,844 in 2010 is reflected in the RCCE curves as shown in Figure 35. Similar curves for DFW are shown in Figure 36.

The RCCEs used in the study to describe the airports operational limits were estimated from past airport utilization data. These estimates provided lower bounds for theoretical airport RCCEs. Actual airport RCCEs are expected to be more inclusive and larger. The airport configuration benefits were first computed using 2010 airport utilization data, then with the higher 2001 utilization data. The 2010 operational limits in 2010 are lower than in 2001 and therefore significantly lower than representative RCCEs. The difference in airport configuration benefits due to 2001 and 2010 RCCEs is instructive as it provides insight on the performance of the optimization algorithm relative to the RCCE changes. Therefore both results are included in the report. No clear relationship can be drawn between changes in RCCEs and expected benefit performance. However, an increased capacity due to larger RCCE naturally accommodates more operations and generally decreases the need to optimize.

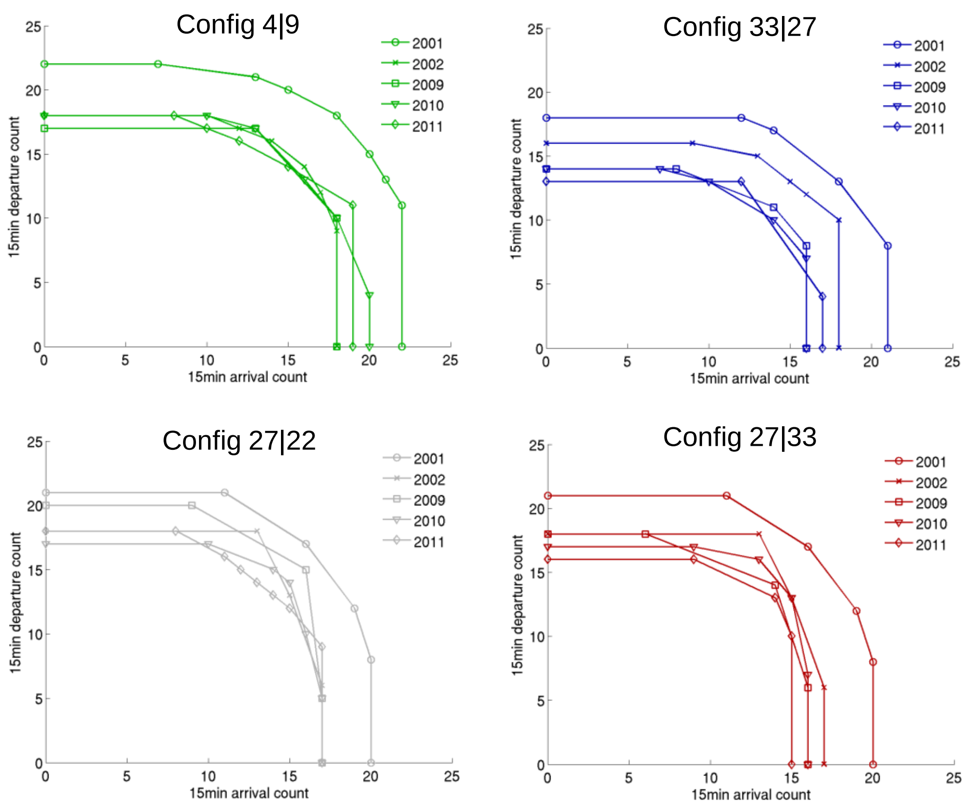


Figure 35: Runway Capacity Envelopes for Boston's Top 4 Configurations for 2001, 2002, 2009, 2010, and 2011. Capacity envelopes in 2001 are consistently larger

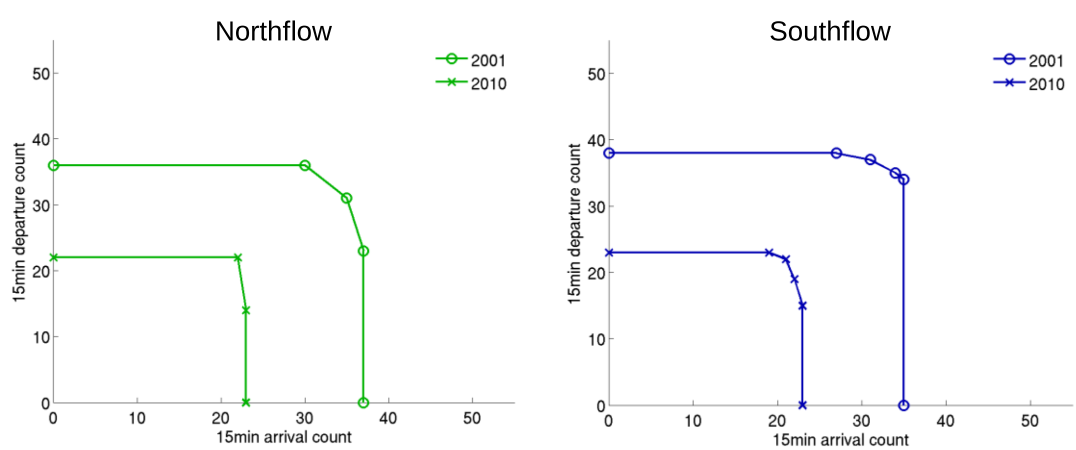


Figure 36: Runway Capacity Envelopes for Dallas/Fort Worth's Top 2 Configurations for 2001 and 2010

Capacity envelopes characterize underlying configurations. However, many configurations are related by added or subtracted runways. Reported configurations can therefore be clustered into groups while maintaining a balance between sufficiently discriminating configurations and excessive number of configurations. FAA ASPM data for Boston shows usage of 67 different configurations during the year 2010. Many of these configurations are inter-related and the exact demarcation between configuration is unclear. Configurations were initially clustered into inclusive supersets resulting in 20 configuration groups, then further reduced to five basic subgroups after consultation with Boston tower controllers.

5.2.4 Feasible Configurations

The effectiveness of the optimization algorithm depends on the judicial selection of feasible configurations. At each time interval, a configuration is chosen from the set of feasible configurations consistent with current meteorological and demand conditions in such a manner as to accommodate the most number of operations. An evaluation of the many factors influencing the feasible set selection at Logan airport (wind direction and seasonal correlations shown in Figure 37 and Figure 38) concluded that the overriding contributing factors to configuration selection are weather and demand. This observation was confirmed by controllers at Logan, JFK, LaGuardia and Newark. The influence of demand on configuration selection is embedded in the capacity envelopes curves. Other factors including noise abatement schedules, runway repair and maintenance are omitted in this study but could be explicitly considered in follow-on work.

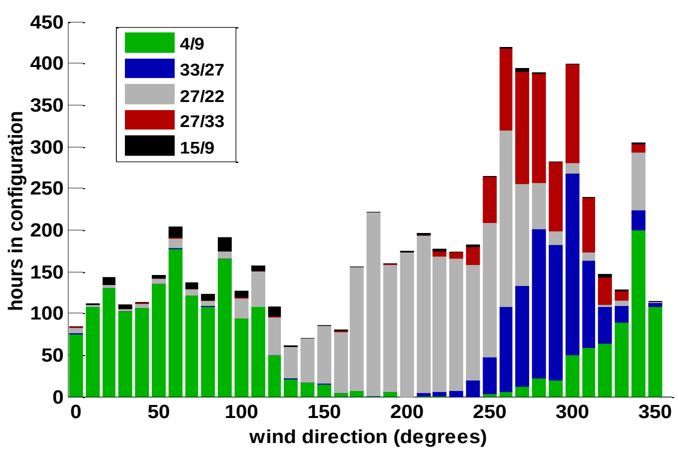


Figure 37: Total Hours in Configuration per Wind Direction at Boston in 2010. The cluster of gray (27/22) in the middle is an example of high usage of configuration over a wind direction range

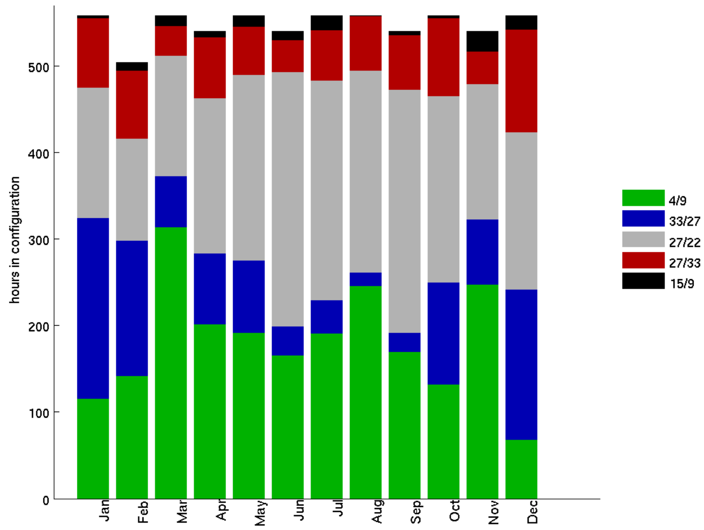


Figure 38: Seasonal Configuration Usage at Logan in 2010. Usage of configuration 27/22 (gray) and 33/27 (blue) are increased during summer and winter months, respectively

The inclusion of a configuration into a feasible set is based on a maximum tailwind of 8 kts and a maximum crosswind of 25 kts. This rule was shown to be 99% inclusive at Boston and DFW over the recorded configurations data for 2010.

5.2.5 The "Binning" Effect and Performance Metric

The RCCEs are constructed from data counted within fixed intervals of time, or bins. The optimization algorithm therefore reports *unserved demand* based on these bins. Actual traffic delays, however are not subject to the artificial bin partitioning and therefore require the conversion of unserved demand (or binned counts) to delays. A statistical conversation factor is derived by measuring actual delays and counting the average number of flights falling within or outside the given bins. Results from this finding are summarized in Table 17 which shows a reasonable agreement between bin size and actual delays, and consistent over bin sizes. Figure 39 shows the estimated delay time from the actual unserved demand in that bin. Bin size zero (no bin) reflects the actual reported delays. The estimated delays for bin sizes other than zero are calculated by multiplying the counted unserved demand by the average minutes per unserved. It is therefore concluded that the bin size itself has negligible impact (at the scale of study) on the conversion to actual delays and counting the unserved is a reasonable reflection of delays.

Table 17: Average Traffic Delays (minutes) Outside Fixed Bin Size Intervals

Bin Size (mins)	3	6	9	12	15	18	21	24	27	30
Arr-min	2.92	5.88	8.78	11.69	13.57	17.63	20.46	23.59	26.35	30.14
Dep-min	2.90	5.80	8.68	11.60	14.16	17.3	20.3	23.32	26.08	28.48

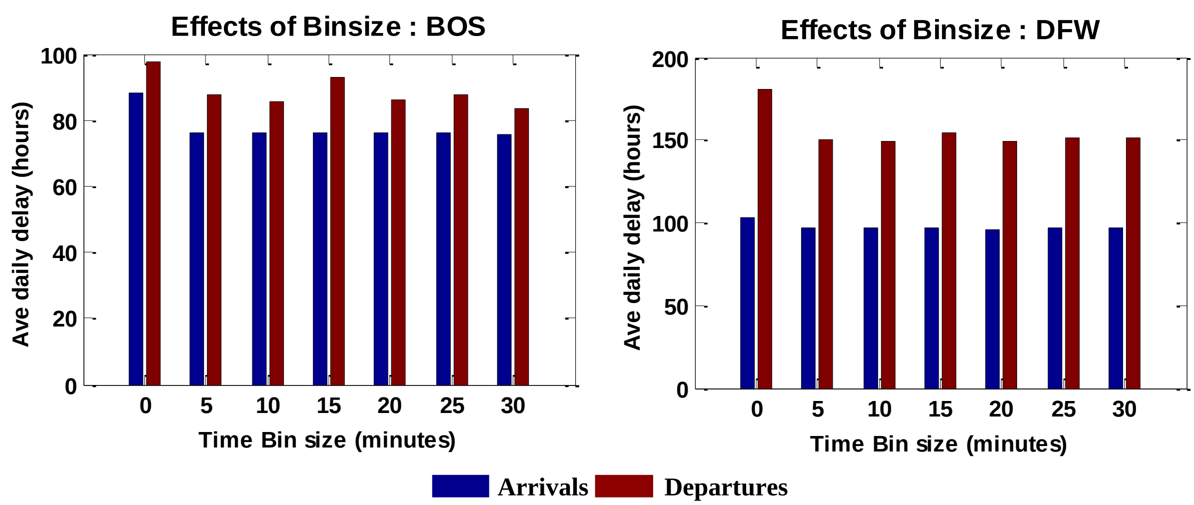


Figure 39: Average Delay of Unserved Traffic Given a Bin Size

5.3 AIRPORT SCOPE

The study primarily focuses on Boston but also includes analysis of Dallas/Fort Worth airport. These airports were selected for their relative isolation to neighboring airports and the research team's extensive prior knowledge. Dallas Fort Worth airport differs from Boston in its higher demand and lower variability in runway directions. DFW's fewer unique runways directions translate into fewer unique configuration options. Analysis to a broader set of airports would be possible to support future benefits assessment processes if desired by FAA.

5.4 RESULTS

The lower performance bounds reported below reflect the imposition of a 15 minute down time (penalty to both arrivals and departures) for every configuration change. The upper performance bounds are the result of not assigning a penalty for configuration changes. The 15 minute down time exaggerates the real operations considering that most configuration changes are coordinated well in advance. Preference was given to arrivals over departures throughout the study.

Results for future years are computed by assuming meteorological conditions similar to 2010 and using the highest capacity 2001 RCCEs observed in past years. The future benefits at BOS and DFW are shown in Figure 40 and Table 18.

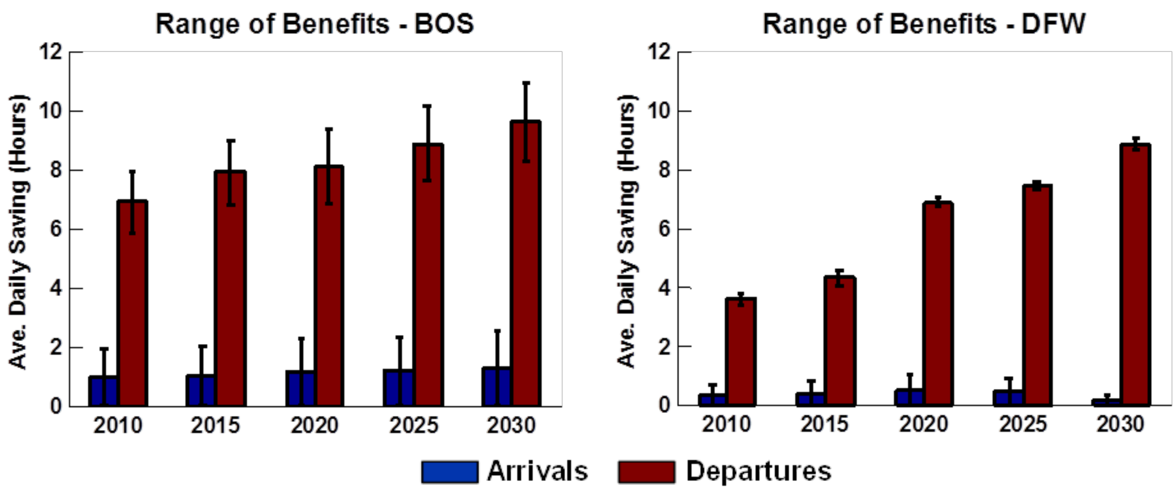


Figure 40: Estimated Range of Benefits of Optimizing Airport Configuration Usage at BOS and DFW Airports

Table 18: Airport Configuration Optimization Results

Airport	Total Delay Savings (khrs)	Fuel Savings (million gallons)	Fuel Savings ($millions)
BOS	76.9	16.9	41.1
DFW	59.4	15.3	37.1

The BOS aggregate delay savings 2015 to 2035 are estimated to be 77,000 hours (68,000 hours for departures and 9,000 hours for arrivals), while for DFW the total estimate is 60,000 hours (57,000 hours for departures and 3,000 hours for arrivals). These translate into fuel savings of 17 million gallons ($41 million) for BOS and 15 million gallons ($37 million) at DFW using the same fuel burn rates and costs as described in the analysis of the other capabilities.

The savings are accentuated with the cost attributed to a configuration change. The greater the configuration change delay costs, the greater the potential benefits. The specific configuration change delays vary with traffic, runway configuration and airspace. The potential benefits to optimizing configuration change in Boston would have led to an aggregate 4–8 hour daily delay savings.

As shown on Figure 40, departure delays are naturally higher than arrival delays due to the normal priority given to arrivals. The ranges of benefits in Boston are greater than that of DFW due to the higher configuration choices in BOS. The benefits at BOS are nearly constant over the years, while DFW sees an increase in benefits. This difference is perhaps due to BOS operating at closer to capacity than DFW.

5.5 DISCUSSION

This study is based a configuration selection model driven by wind and traffic demand. Wind data (METAR) provides the inclusion rules for constructing the feasible configuration sets while traffic demand constraints are summarized within the capacity envelopes. Excluding demand (and other operational factors) from the feasible set selection process provides the optimization algorithm with greater choice of configurations to select from and therefore yields an upper bound on benefits. It is therefore suggested that a more precise follow up study include demand in the feasible configuration set selection. In addition, wind forecasts are rarely accurate for long periods ahead as assumed in this study. One can therefore assume a degradation in performance when the configuration sequence is predicted far in advance. It is recommended that a more realistic approach to wind forecast errors be part of a more comprehensive future work.

6. SAFETY ASSESSMENT

6.1 BACKGROUND

Increases in traffic volume and complexity will increase controller workload and potentially decrease situational awareness for managing traffic in the airport environment. Given that TFDM aims to enable safe and efficient operations under NextGen, it is critical to analyze potential safety impacts and determine what types of real-world safety issues can be prevented or mitigated by TFDM. With this goal in mind, MIT LL conducted a data-driven safety assessment involving a comprehensive review of aviation accident and incident databases. Similar assessments have been conducted for other aviation systems such as Runway Status Lights (RWSL) [Wilhelmsen (1994)].

Most towered airports will receive the Flight Data Manager (FDM) element of TFDM to support the electronic distribution and tracking of flight data and clearances, supporting situation awareness and reducing workload associated with maintaining an accurate picture of the traffic situation in increasingly complex circumstances. Airports with surface surveillance capabilities (i.e., ASDE-X) will also receive the enhanced surveillance display (the Tower Information Display System, or TIDS) and therefore benefit from enhanced processing of surveillance data that enables both intent and state-based conflict detection. Finally, the decision support tools which have been the focus of the other analyses presented in this document will be introduced at key airports that will allow controllers to monitor and manage traffic more effectively and will provide advance notice of hazardous situations. Realizing these safety benefits will reduce the likelihood of accidents and incidents, providing both qualitative and quantitative evidence to support the implementation of TFDM. From a qualitative perspective, it is critical to consider accident precursors and the likelihood that safety will be compromised as complexity increases. Reason's Swiss cheese model of accident causation [Reason (1990)] provides a framework for systematically analyzing accident precursors across four different layers, as shown in Figure 41.

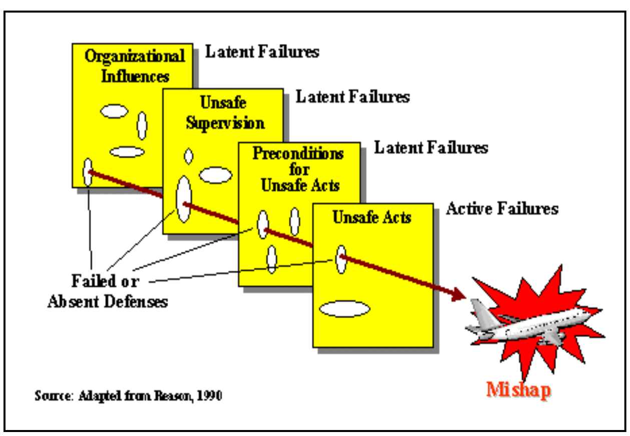


Figure 41: The Reason Model and Accident Causal Chain

The Human Factors Analysis and Classification System (HFACS) relies on this model to identify the causal sequence of events during accident investigations [Wiegmann & Shappell (2003)]. It is often the case that accidents and incidents involve multiple failures lining up across various layers due to failed or absent defenses. Within an air traffic control context, latent failures in the "organizational influences" layer may involve inappropriate processes or a climate conducive to complacency. While TFDM does not directly address organizational influences, this layer is critical in that it can impact performance at all other levels. The next layer refers to "unsafe supervision" and captures strategic issues such as planned inappropriate actions (e.g., maintaining or choosing an airport configuration not aligned with environmental constraints). Decision support tools provided through the TFDM Supervisor Display provide support for this layer. Moving to the next layer in the model, "preconditions for unsafe acts" includes both environmental (e.g., reduced visibility) and operator state (e.g., high workload, low situational awareness) factors. TFDM provides defenses at this layer through improved surveillance and the consolidation of stove-piped systems, allowing easier access to information. The final opportunity for accident prevention is captured by the "unsafe acts" layer where errors or violations may take place. Decision errors (e.g., decision to issue takeoff clearance while another aircraft is landing) and perceptual errors (e.g., misjudging aircraft location) occur at this layer and are targeted by many aspects of TFDM. Electronic flight data, for example, tracks aircraft state and provides earlier alerting to potentially hazardous conditions (e.g., runway incursions). In addition, decision support tools provide tactical support for monitoring the airport surface and alerting the controller to situations in need of attention (e.g., taxi non-conformance). In summary, TFDM addresses issues in multiple layers of the accident causal chain, thereby targeting the root causes of accidents. Existing tools (e.g., AMASS, RWSL) only target the mishap itself.

6.2 ANALYSIS METHODOLOGY

The safety analysis followed a structured approach in identifying current and future estimates of TFDM safety-relevant benefits, as shown in Figure 42. For current year estimates, MIT LL utilized archived accident and incident data to determine observed frequencies of safety events. Baseline safety impacts were generated utilizing ATO-F monetization values [FAA/ATO (2011)] and compared to the safety impacts associated with implementing TFDM capabilities in order to estimate the theoretical monetized benefit of TFDM within the current year. For future year estimates, MIT LL utilized the demand and capacity forecasts (as in the other areas described) in combination with the observed safety event frequencies to build a safety event forecast model that predicted future frequencies of safety events in five-year increments. This process was informed by future year safety models employed for other programs such as the ASDE-X system [Johnson (2005)]. Future baseline safety monetized impacts were compared to the future monetized impacts when TFDM capabilities were implemented in order to estimate the safety benefits of TFDM for 2015, 2020, 2025, and 2030. The key output metrics of this analysis included personal injury and aircraft damage monetized impact savings.

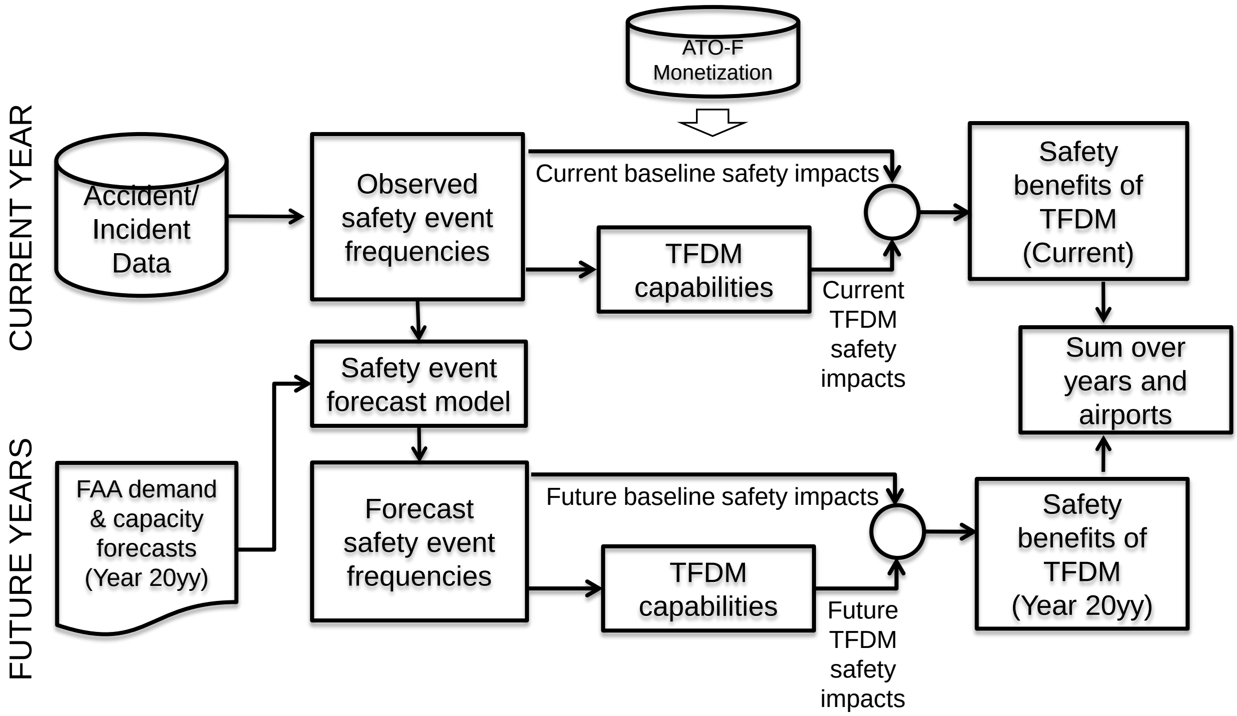


Figure 42: Safety Assessment Analysis Methodology

Several databases maintained by the FAA through the Aviation Safety Information Analysis and Sharing (ASIAS) System were utilized to ensure comprehensive coverage of safety-related events that could potentially be prevented or mitigated by TFDM. Each relevant database is described below.

The National Transportation Safety Board (NTSB) maintains a database of civil aviation accident and selected incident investigation reports. Example analyses conducted on this database include an examination of situation awareness errors [Endsley (1995)] as well as weather-related accidents [FAA/Weather (2010)]. NTSB reports represent thorough investigations of events associated with the operation of an aircraft where any person suffers death or serious injury or any aircraft receives substantial damage (i.e., damage or failure that negatively affects an aircraft's structural strength, performance, or flying characteristics, and which would require significant repair or replacement of the affected component or system; NTSB Part 830). The NTSB divides accidents into four categories:

1. Major – an accident in which an aircraft was destroyed, there were multiple fatalities, or there was one fatality and an aircraft was substantially damaged.
2. Serious – an accident in which there was either one fatality without substantial aircraft damage, or there was at least one serious injury and an aircraft was substantially damaged.

3. Injury – a nonfatal accident with at least one serious injury and without substantial aircraft damage.
4. Damage – an accident in which no person was killed or seriously injured, but in which any aircraft was substantially damaged.

Only completed investigations were included in this analysis to ensure that probable causes had been identified.

The FAA Accident/Incident Data System (AIDS) database is utilized to record civil aviation incident data not accounted for by the NTSB because they do not meet the NTSB-defined aircraft damage or personal injury thresholds described above, but do or could affect the safety of operations (NTSB Part 830). This database primarily captures safety-related events such as flight deck caution messages, warning lights, engine fires, flaps failures, etc. that are not impacted by air traffic control operations. An initial review of a year's worth of reports from this database confirmed that it is not relevant to the TFDM safety analysis.

The Aviation Safety Reporting System (ASRS) is a voluntary reporting system utilized by pilots, air traffic controllers, and other aviation industry personnel to submit subjective accounts about safety-related aviation events. Previous analyses utilizing this database have focused on flight deck automation issues [Funk et al. (1999)], memory errors [Nowinski et al. (2003)], and the effect of interruptions [Damos & Tabachnick (2001)], to name a few. The ASRS database is subject to a number of strengths and weaknesses given its voluntary nature. With respect to strengths: (1) incident participants generally provide a detailed account of the reported event, (2) a large number of incidents are reported, and (3) the reports can be viewed as ecologically valid given that the incidents occurred in an operational environment [Jones & Endsley (1995)]. On the other hand, limitations of the database include: (1) reports cannot be viewed as a random sample of all incidents, (2) incident accounts are subject to possible reporter bias, and (3) incidents are not investigated or verified through independent means [Chappell (1994), Degani et al. (1991)]. Despite these limitations, reported incidents provide valuable qualitative information regarding the types of hazards, accident precursors, and safety-related issues that could potentially be prevented or mitigated by TFDM. Only reports submitted by tower air traffic controllers were included in this analysis.

The FAA Operational Error/Deviation System (OEDS) is the official source of air traffic control operational error data. Example analyses utilizing this database include the relationship of controller age to en route operational errors [Broach & Schroeder (2005)] and the relationships between error occurrence, controller workload, and causal factors involved [Rodgers (1993)]. The database includes information regarding the number of aircraft under controller responsibility at the time of the error, whether or not training was in progress, time on position, etc. This database is not publically available and access could not be obtained for this IID analysis; MIT LL requests that FAA and MCR provide access to these data for the Final Investment Decision if possible.

MIT LL reviewed Part 121 accidents/incidents that occurred over a five-year period between January 2005 and December 2009. We focused on Part 121 operations as they represent scheduled commercial air carriers generally operating out of controlled airports. Only final reports were utilized to ensure more accurate reporting of causal factors. These selection criteria produced a total event count of 560 (NTSB: 247; ASRS: 313).

A coding spreadsheet was developed to collect relevant data from each of the selected reports. Information gathered on each report included the following:

- Source (NTSB, ASRS);
- Report ID;
- Date (month, year);
- Time of day (0001–0600, 0601–1200, 1201–1800, 1801–2400);
- Airport;
- Position (Ground Control, Local Control, Local Assist, Supervisor);
- Flight phase (Taxi, Takeoff, Climb, Approach, Landing);
- Weather conditions (VMC, IMC, marginal, not reported);
- Injury category (e.g., minor, serious, fatality);
- Aircraft damage (none, minor, substantial, destroyed);
- Contributing factor(s) (i.e., HFACS code);
- Narrative URL;
- Synopsis;
- Outcome (e.g., runway incursion); and
- Effectiveness ratings of TFDM components.

Contributing factors and effectiveness ratings of TFDM components were inferred by the analyst given information provided in the individual reports. HFACS codes, presented hierarchically in Figure 43, were utilized to systematically categorize relevant contributing factors. Definitions for each code are provided in Appendix C [DoD HFACS (2005); Wiegmann and Shappell (2001)].

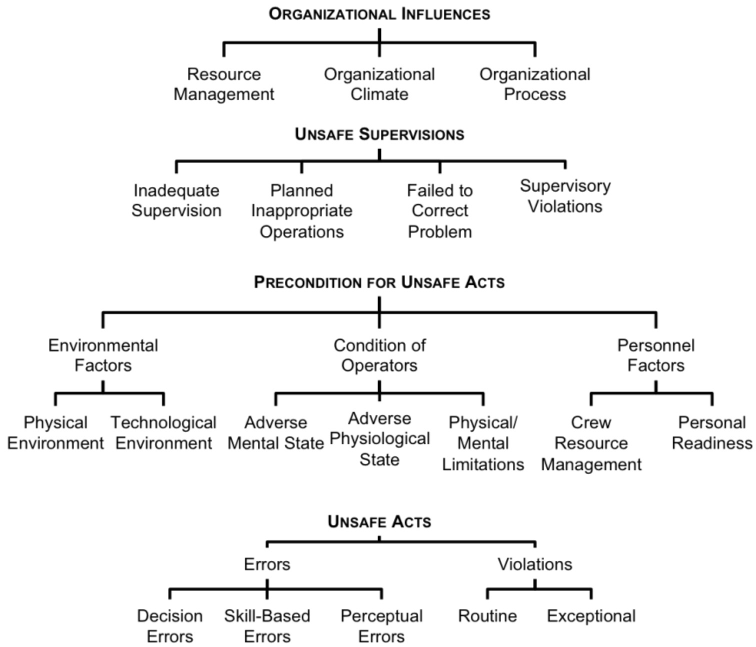


Figure 43: HFACS Codes

A systematic method for generating probabilistic estimates of benefits for a technology not yet deployed [Barnett and Paull (2004)] was utilized to produce effectiveness ratings for TFDM components. While this method generally involves a panel of experts to provide effectiveness ratings, no such panel was available to support this analysis (but could be explored in the future). Instead, an aviation human factors expert with piloting experience rated the likelihood that individual safety-related events could have been mitigated or prevented by TFDM components. Input was sought from FAA and MCR to ensure no double-counting of benefits from other systems were included in the TFDM safety analysis. The TFDM components were considered incrementally according to planned implementation phasing; namely, consolidation and electronic flight data were rated for all airports, improved surveillance was rated for all ASDE-X airports, and then decision support tools were also rated for all ASDE-X airports. Specifically, the rater considered three questions for each analyzed incident:

- Would the availability of consolidated/integrated systems and electronic flight data have prevented the event?
- Would the availability of consolidated/integrated systems and electronic flight data plus improved surveillance have prevented the event?
- Would the availability of consolidated/integrated systems and electronic flight data plus improved surveillance and decision support tools have prevented the event?

Responses to these questions were provided along a five-point scale ranging from "almost definitely no" to "almost definitely yes" with intermediate responses of "probably no," "50/50," and "probably yes." These responses were translated into probabilies as follows:

- Almost Definitely No 0%
- Probably No 25%
- 50/50 50%
- Probably Yes 75%
- Almost Definitely Yes 100%

This method allowed for calculations of incremental effectiveness per TFDM component as well as an aggregate effectiveness rating of the TFDM system as a whole.

6.3 AIRPORT SCOPE

Following a thorough review of each accident report meeting the selection criteria defined previously, a subset of the reports were found to be relevant to the TFDM safety analysis. 115 of the reports (24 from the NTSB database and 91 from ASRS) originated from the TFDM analysis airport set shown in Appendix A and these were used to define a "lower bound" on the safety benefits assessment monetization. 129 of the full set of reports (25 from the NTSB database and 104 from ASRS) were considered relevant to TFDM and originated from *all* airports captured by these databases. Consideration of this larger set of airports (in addition to the TFDM analysis airport set) was recommended by safety expert Prof. Arnold Barnett at MIT to increase the sample size and to include major incidents (in particular the Lexington airport accident from 2006) which could theoretically occur at any airport. This larger airport set produced an "upper bound" on the safety benefits assessment monetization. Results from both airport sets are presented in the following section.

6.4 RESULTS

The following five charts present contextual information summarizing the analyzed incidents by airport, tower controller position, weather conditions, phase of flight, and contributing factor for the "lower bound" airports. While these results do not directly support the monetization of safety benefits, it is useful to examine the circumstances under which historical incidents occurred to provide a better understanding of accident precursors.

Forty airports were represented in the reports deemed relevant to the TFDM safety analysis, 30 of which are classified as ASDE-X airports. A breakdown of the number of safety events per airport (both within and across databases) is shown in Figure 44. The median number of analyzed events per airport was 1.5, with a range from one to 14. Note that four ASRS reports did not indicate at which airport the event occurred.

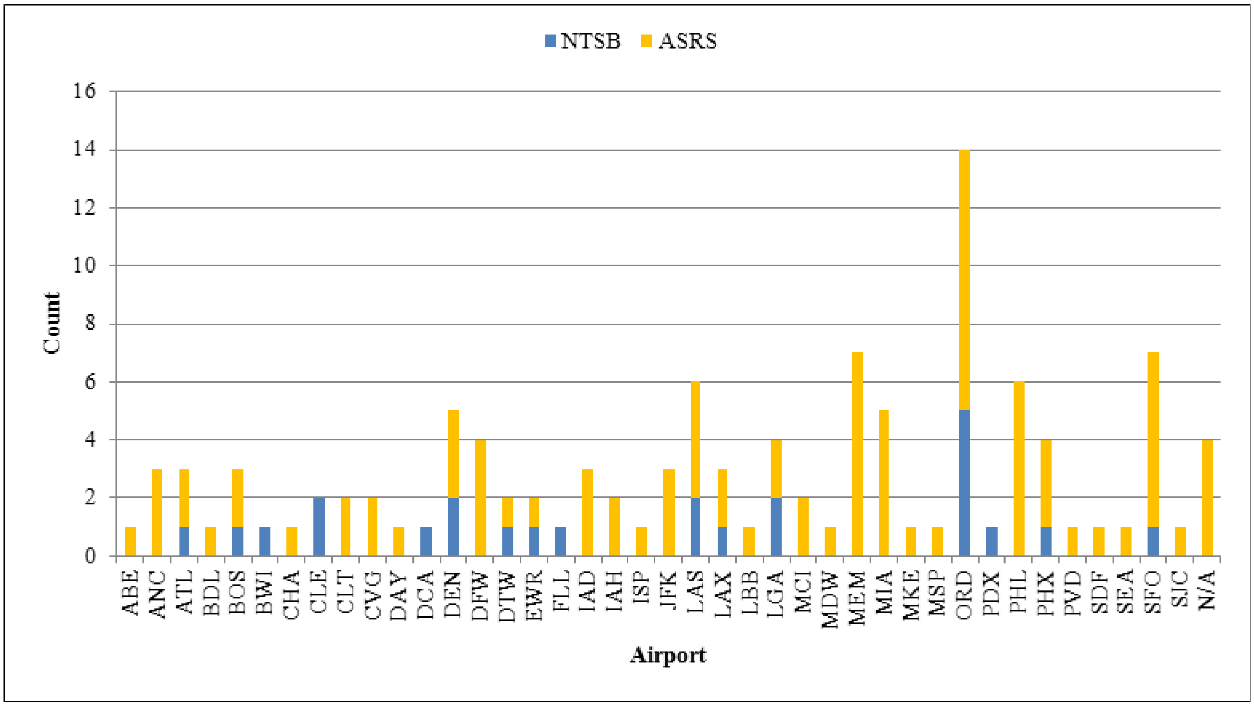


Figure 44: Lower-Bound Airport Summary Results by Database

As shown in Figure 45, an overwhelming majority of analyzed events included a local controller (note that incidents may involve more than one controller position). The local controller is responsible for the active runway surfaces, clearing aircraft for takeoff or landing and ensuring that prescribed runway separation exists at all times. Although the time spent taking off and landing represents a small portion (~6%) of the total time spent in flight, over half of all accidents occur during the takeoff and landing phases [Boeing (2011)]. In looking at the breakdown of contributing factors specific to the local control position, the vast majority of cases involved adverse mental states (e.g., high workload; 47) and decision errors (e.g., inappropriate takeoff clearance; 39), reflecting the complexity involved with operations during the takeoff and landing phases. Other key contributing factors to analyzed incidents involving a local controller include the technological environment (e.g., system failures; 26), skill-based errors (e.g., visual scanning disruptions; 23), and the physical environment (e.g., inclement weather; 22).

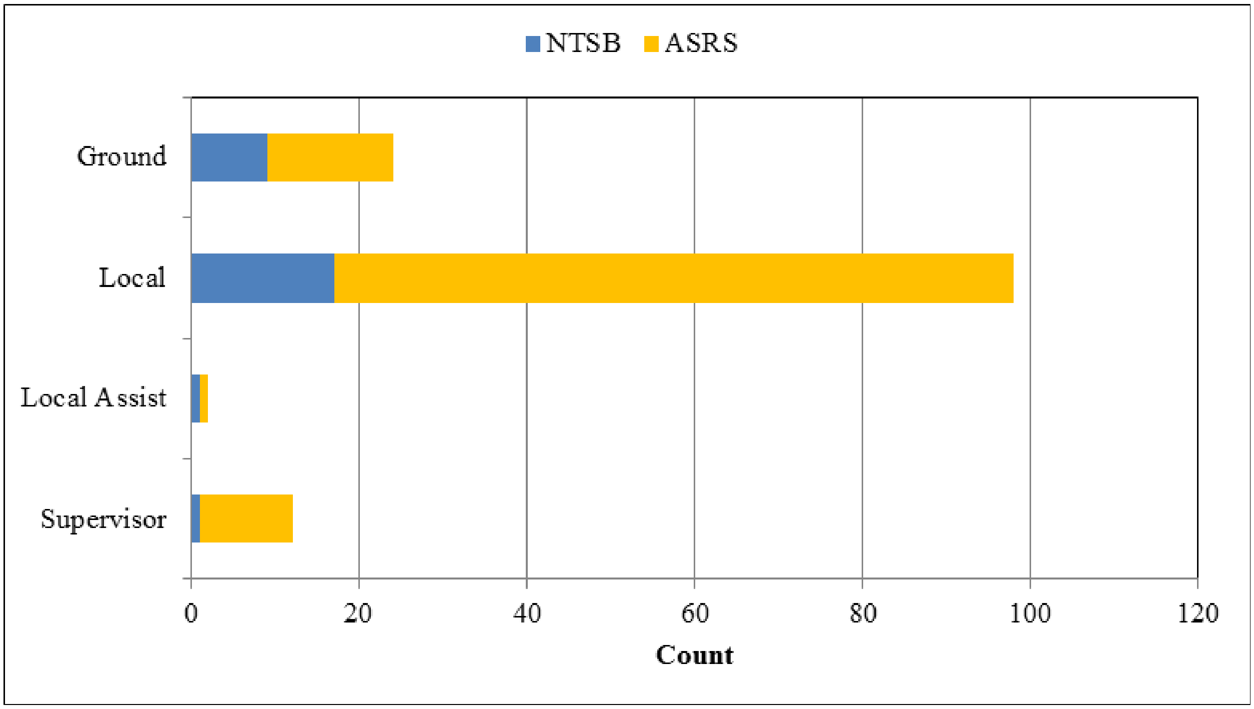


Figure 45: Tower Position Summary Results by Database

The majority of analyzed safety events occurred during visual meteorological conditions (VMC). Figure 46 shows the breakdown of reports by VMC, instrument meteorological conditions (IMC), and marginal weather conditions (note that there were five cases in the ASRS database in which the weather was not reported). In interpreting these findings, it is important to keep in mind the percentage of time spent in VMC versus IMC at any given airport. For example, the FAA reported that in 2004, ATL spent 73% of the time in VMC, 16% of the time in marginal VMC, and 11% of the time in IMC [Kang et al. (2007)]. Assuming similar distributions at other airports, it would be expected that a larger number of safety-related events would occur under VMC than IMC simply due to the amount of time spent under these conditions. With respect to contributing factors, adverse mental states (49) and decision errors (41) account for the vast majority of safety-related events during VMC, while the physical environment (8) was the biggest contributing factor during IMC.

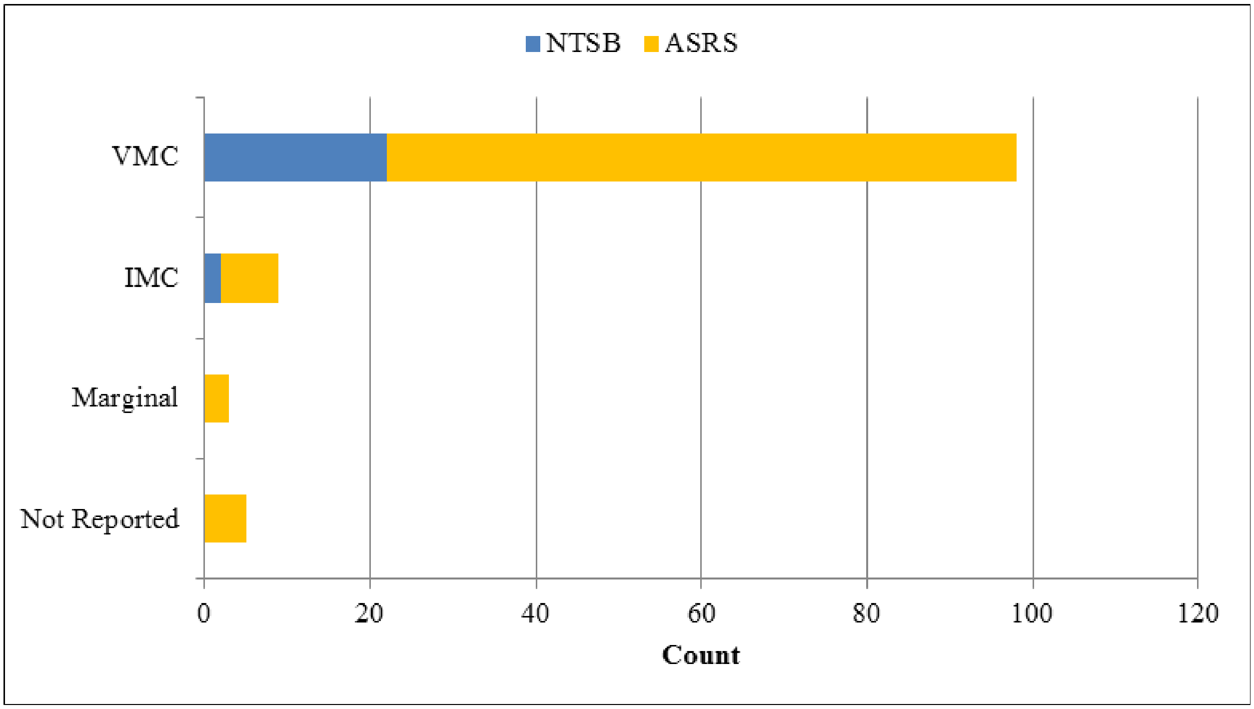


Figure 46: Weather Condition Summary Results by Database

Analyzed incidents represent the full range of tower operations in terms of phase of flight. It is clear from Figure 47 that the majority of safety events involved at least one aircraft that was taking off (note that events involving multiple aircraft may represent more than one phase of flight). As mentioned previously, over half of all accidents occur during the takeoff and landing phases. Analyzed incidents represent the full range of tower operations in terms of phase of flight. It is clear from Figure 47 that the majority of safety events involved at least one aircraft that was taking off (note that events involving multiple aircraft may represent more than one phase of flight). As mentioned previously, over half of all accidents occur during the takeoff and landing phases. Interestingly, more safety-related events occurred during the taxi than landing phase across our analyzed incidents. Adverse mental states (44) and decision errors (43) are the leading contributing factors during the takeoff phase, while adverse mental states are most implicated during the taxi (32) and landing (17) phases. There are no clear trends for contributing factors during the climb or approach phases of flight.

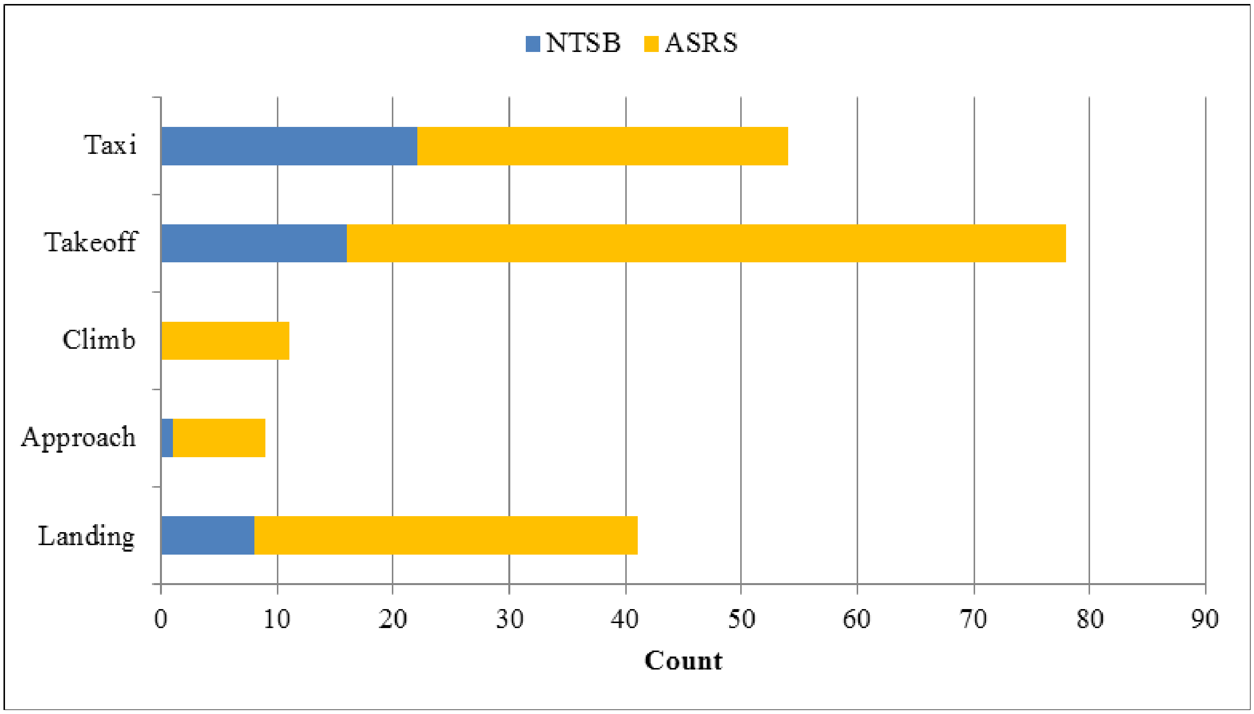


Figure 47: Flight Phase Summary Results by Database

As discussed previously, contributing factors to analyzed incidents were classified according to HFACS codes, the results of which are shown in Figure 48. Nine of the 19 HFACS codes were identified as relevant to the TFDM safety assessment. All safety events were associated with at least one HFACS code, with a range from 1 to 4 HFACS codes per incident (mean = 1.8). There appear to be three natural groupings of contributing factors in terms of their frequencies in contributing to analyzed incidents. Decision errors and adverse mental states comprised the majority of coded safety-related events. Skill-based errors and both the physical and technological environments form the second grouping. Perceptual errors, adverse physiological states, crew resource management, and planned inappropriate operations form the final grouping. From a human factors perspective, the breakdowns across these contributing factors could be utilized to drive design requirements in future systems to ensure that proposed solutions (technological or otherwise) actually address existing safety threats. For example, the high frequency of decision errors contributing to safety-related events points to the need for improved decision support systems within the air traffic control tower.

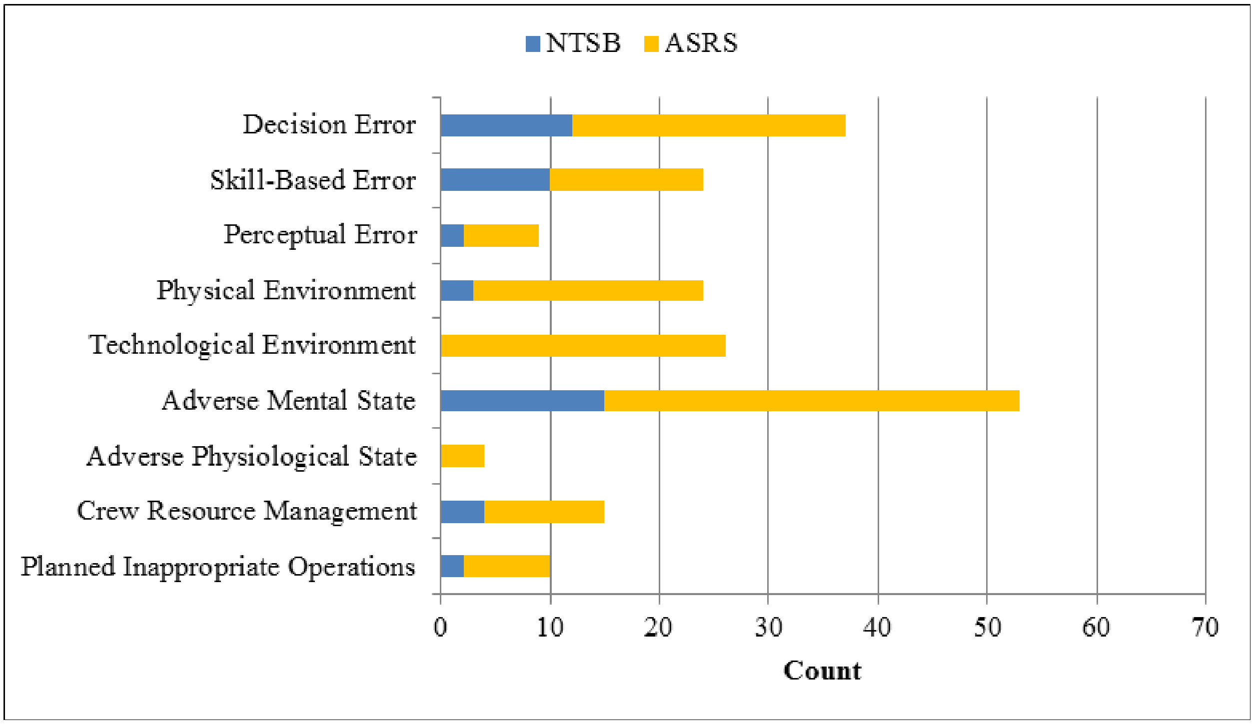


Figure 48: Contributing Factors Results by Database

Of most relevance to the actual monetization of TFDM safety benefits are the TFDM effectiveness ratings. Table 19 summarizes the aggregate effectiveness results (averaged over all NTSB and ASRS accidents/incidents considered) according to incremental benefits provided by a phased implementation of TFDM components. TFDM core implementation involving the consolidation of systems as well as the availability of electronic flight data has an effectiveness of 27% in preventing or mitigating analyzed incidents. An example incident that could have been prevented or mitigated by electronic flight data is the Boston 2005 runway incursion (NTSB event ID 20050624X00863) in which two aircraft (EIN132, USA1170) were cleared for takeoff on intersecting runways within five seconds of one another; the FDM would have alerted the controller when the second aircraft (USA1170) was cleared for takeoff and the controller would have been able to immediately cancel the takeoff clearance. Adding improved surveillance to the cases involving ASDE-X airports provides an average incremental effectiveness of 40%. Enhanced conflict detection enabled by this capability could have prevented or mitigated the SFO 2007 runway incursion (NTSB event ID 20070610X00701) in which the controller forgot about a landing aircraft (SKW5741) and cleared another aircraft (RPA4912) for takeoff on an intersecting runway. TFDM would have alerted the controller when RPA4912 was cleared for takeoff given surveilled information indicating that SKW5741 was crossing the landing threshold of the intersecting runway. Finally, adding decision support tools to those same cases involving ASDE-X airports provides an average incremental effectiveness of 12%. An example incident that would have been prevented by a decision support tool is the Denver 2007 runway incursion (NTSB event ID 20070110X00037) in which an aircraft (LYM4216) missed its taxiway turn due to inclement weather and ended up turning onto an active runway on which another aircraft (FFT297) was attempting to land. Taxi conformance monitoring, specifically, would have alerted the controller to this situation when LYM4216 missed its intended taxiway. In total, full implementation of TFDM reveals an effectiveness of 79% in preventing or mitigating safety-related events across analyzed incidents.

Table 19: TFDM Effectiveness Ratings

TFDM Component	Effectiveness Rating	
	Lower Bound Airport Set	**Upper Bound Airport Set**
Consolidated/integrated systems and electronic flight data	27%	25%
Plus improved surveillance	40% increment	42% increment
Plus decision support tools	12% increment	15% increment
Total Effectiveness Rating	79%	82%

It is important to note that the above values were estimated based on incidents occurring at specific airports in conjunction with current plans for implementing ASDE-X at a select number of airports which, as previously discussed, defined the “lower bound” airport set. This assumption limited the incremental effectiveness ratings of improved surveillance and decision support tools to the relatively small number of ASDE-X airports. Assuming that the types of incidents that occur at non-ASDE-X airports could theoretically also occur at ASDE-X airports, however, presents an argument for uncoupling specific incidents from the actual airports at which they occurred. Taking this perspective and reassessing the effectiveness of TFDM components in preventing or mitigating analyzed incidents reveals a slightly different picture, namely, minor increases in the incremental effectiveness ratings of improved surveillance and decision support tools, ultimately leading to a higher overall effectiveness rating (82%) of TFDM as a whole. These “upper bound” TFDM effectiveness ratings are presented on the right side of Table 19.

FAA-recommended FY11 economic values for aircraft damage and personal injuries (see Table 20; from the March 2011 “Economic Information for Investment Analysis” data package [FAA/ATO (2011)]) were used to monetize the average impacts associated with accidents/incidents meeting NTSB-defined thresholds. Many incidents did not meet NTSB-defined aircraft damage or personal injury thresholds and therefore could not be monetized according to the values in Table 20. Also, because some events represent potential accident or incident precursors (as opposed to actual accidents) there were no direct costs to fatalities, injuries, or aircraft damage.

Table 20: FAA-Recommended FY11 Economic Impacts by Accident/Injury Category

Category	Economic Impacts
Replacement aircraft cost	$13,650,000
Restoration aircraft cost	$4,410,000
Minor repairs aircraft cost	$500,000 (MIT estimate)
Fatality	$6,000,000
Critical injury	$4,575,000
Severe injury	$1,125,000
Serious injury	$345,000
Moderate injury	$93,000
Minor injury	$12,000

Table 21 presents the safety-related costs associated with relevant monetizable accidents, as well as the effectiveness ratings for various TFDM capabilities assessed using the method previously described. Note that only a small subset of the 100+ incidents deemed relevant to TFDM could be monetized using the FAA-recommended personal injury and aircraft damage criteria. The lower bound costs do not include the Lexington (LEX 2006) accident given that airport is not projected to have ASDE-X and would therefore not benefit from improved surveillance and/or decision support tools under the conservative assumption utilized in calculating the lower bound economic benefits. Total safety costs therefore total a lower bound of $42.8 million. This additional monetizable accident was included in the upper bound total safety costs of $350.8 million (i.e., nearly a factor ten higher than the lower bound). The difference between the upper and lower bound safety costs arise directly from the accident at LEX which involved 49 fatalities, one serious injury and a destroyed aircraft for a total cost of $308 million.

These raw safety costs were then weighted by the effectiveness ratings shown in Table 21 to determine the potential monetizable safety benefits of different capabilities of TFDM. The integration and electronic flight data elements of TFDM was not considered to have any potential to mitigate or prevent the incidents given in Table 21 and hence would not have any monetizable safety benefit. However, improved surveillance was considered capable of mitigating or preventing $18.3 million in monetizable safety impact for both the lower and upper bound cases. The addition of TFDM DST capabilities was estimated to lead to $24.9 million and $333 million of monetizable benefit for the lower and upper bound cases respectively for these historical incidents.

Table 21: Economic Values and TFDM Effectiveness Ratings per Accident

Accident/ Incident	Injury Impacts	Aircraft Damage Impacts	Combined Impacts (Lower)	Combined Impacts (Upper)	ER Integ/elec flt data	ER + Imp surv	ER + DSTs
DEN 2008	$2.6m	$4.4m	$7.0m	$7.0m	0%	0%	75%
DCA 2008	---	$4.9m	$4.9m	$4.9m	0%	50%	50%
ORD 2007	---	$4.9m	$4.9m	$4.9m	0%	50%	50%
LAS 2007	---	$0.5m	$0.5m	$0.5m	0%	100%	100%
ATL 2007	---	$4.9m	$4.9m	$4.9m	0%	75%	100%
ORD 2007	---	$4.9m	$4.9m	$4.9m	0%	50%	50%
LGA 2006	---	$4.9m	$4.9m	$4.9m	0%	50%	75%
LEX 2006	$294.3m	$13.7m	---	$308m	0%	0%	100%
PHX 2005	---	$4.9m	$4.9m	$4.9m	0%	50%	50%
PDX 2005	---	$4.9m	$4.9m	$4.9m	0%	50%	50%
EWR 2005	---	$5.4m	$5.4m	$5.4m	0%	50%	50%
	Total impacts		$42.8m	$350.8m			

In order to estimate the potential future monetizable safety benefits of TFDM, the results based on historical incidents were projected into the future accounting for traffic growth. Other aviation system enhancement safety benefits assessments have utilized a (traffic level)^2 relationship between traffic level and safety incidents [Barnett *et al.* (2000), Barnett and Paull (2004)]. This relationship is considered most appropriate for incidents involving two aircraft. Most of the analyzed incidents in this study did involve two aircraft, but the influential LEX accident involved only one aircraft. While other extrapolation techniques have been proposed (e.g., linear, cubic), the implications on the results are relatively small compared to the impacts of choosing the lower bound or upper bound airport set. In fact, Barnett and colleagues point out that (traffic level)^2 may be a conservative estimate given the probability of error due to increased controller workload associated with increased traffic levels over time. Therefore, the (traffic level)^2 relationship is used as the basis for the TFDM safety assessment results according to the equation shown below, but errors bars based on linear and cubic extrapolations are also presented.

$$\$\,TFDM\ safety\ benefit_{Future\ years} = \left(\frac{N_{Future\ years}}{N_{Current\ years}}\right)^2 . \sum_{\substack{historical\\ incidents,\, i}} ER_i . \$_i$$

where N is the traffic level in the air transportation system, ER_i and $\$_i$ are the effectiveness ratings and injury/damage costs of each of the historical incidents i outlined in Table 21. The sum of these $ER_i.\$_i$ products are the $18.3–333 million values given above. Future year traffic levels were determined from the 2010 FAA Terminal Area Forecast (TAF) from ASPM [FAA/ASPM (2011)] using values for total operations at towered airports over five year periods. From this, $N_{Current\ years}$ = 187 million; $N_{2015\text{-}2019}$ = 206 million; $N_{2020\text{-}2024}$ = 223 million; $N_{2025\text{-}2029}$ = 243 million and $N_{2030\text{-}2034}$ = 255 million. The resulting estimates of TFDM monetizable safety benefit in future years using the equation given above are shown in Figure 49 below.

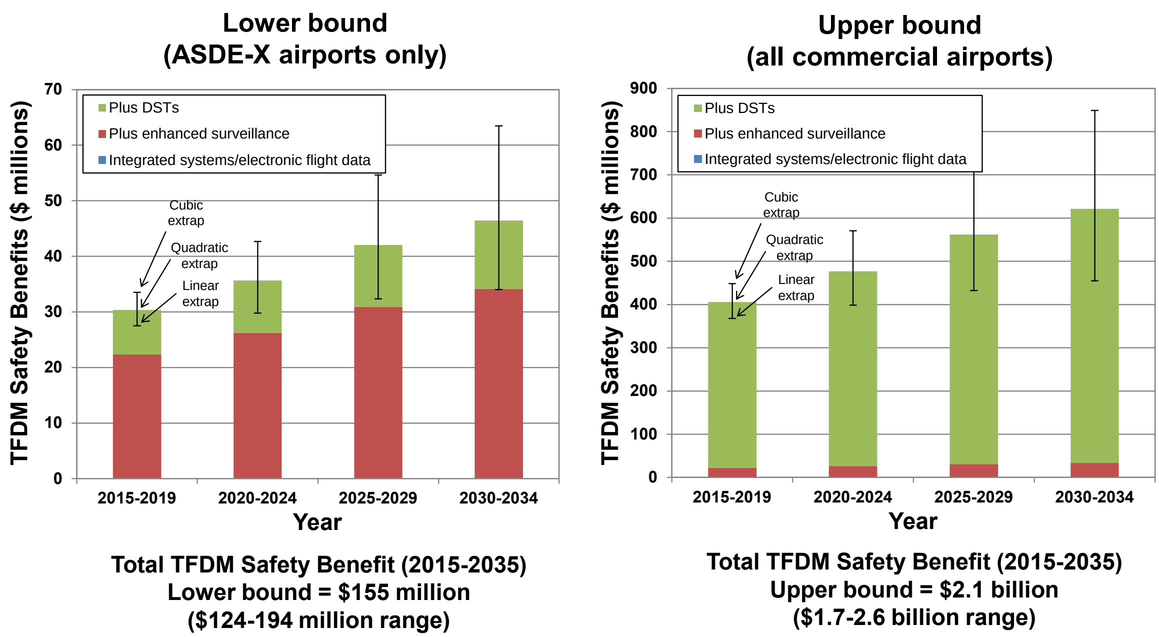


Figure 49: Future Year TFDM Monetizable Safety Benefits

6.5 DISCUSSION

The estimated total TFDM monetizable safety benefit for the 2015–2035 time period is $155 million for the lower bound estimate and $2.1 billion for the upper bound estimate using the (traffic level)^2 extrapolation. The lower bound estimate varies from $124–194 million with the alternate linear and cubic extrapolation methods, while the upper bound varies from $1.7–2.6 billion. Again, the large difference in lower and upper bound estimates is driven by the absence or presence of the LEX accident in the extrapolation method, i.e., whether at least one fatal accident is estimated to be prevented every five years through the deployment of TFDM. The extrapolation method is seen to have a much smaller impact on the monetized results compared to the choice of whether to use the lower or upper bound airport set as the basis for that extrapolation. These results reinforce that there are minimal monetized safety benefit from integrated systems and electronic flight data aspects of TFDM (although user acceptance and efficiency benefits of these aspects might be significant and are recommended to be explored further); enhanced surveillance aspects of TFDM are the major element of the lower bound safety benefit estimate; while DST aspects are the major element of upper bound safety estimate (due to high ER value for DST prevention of the LEX accident).

This page intentionally left blank.

7. PROPOSED FUTURE MIT LL BENEFITS ASSESSMENT ACTIVITIES

This section summarizes the recommended follow-on work which MIT LL believes would be valuable for the TFDM benefits assessment processes. The next phase of the investment process is anticipated to focus on a reduced set of "core" TFDM capabilities including situation awareness and information management functions on the FDM and TIDS primarily for the Ground and Local controller, as opposed to advanced DST capabilities. These non-DST core component benefits were not covered in this analysis except for their impacts on safety discussed in the previous section. Accordingly, it will be important to demonstrate operational user acceptance of the core components as well as data supporting the extent to which those core components improve operational efficiency and safety in stressing traffic environments beyond what is possible in current-day field demonstrations.

For follow-on work, MIT LL recommends a portfolio of benefits assessment activities be pursued, including learning from operational field testing (whenever possible), Human-in-the-Loop (HITL) simulations and computer modeling. Each of these strategies has different strengths and weaknesses in terms of realism, scope and controllability, but together they allow a thorough benefits assessment to be conducted in the context of metrics such as operational performance, environmental performance, user acceptance and safety. In particular, computer modeling allows the theoretical benefits of key TFDM capabilities to be estimated into the future, but operational testing and human and in the loop simulation enable validation and/or calibration of these theoretical benefits to make them more operationally realistic, as well as allowing the collection of user acceptance and other critical human factors data.

Lincoln Laboratory is well placed to continue to support TFDM benefits assessment activities given its role in prior benefits phases, the ability to build upon its existing computer modeling approaches used for the IARD and IID interim analyses, and its sophisticated HITL capabilities. Proposed future tasks are detailed below.

7.1 PROPOSED TASK 1: SIMULATION AND ANALYSIS OF CORE TFDM CAPABILITY BENEFITS

Lincoln Laboratory strongly recommends that HITL studies be pursued for TFDM. This is especially important given the nearer term "core" TFDM deployment will not include many DSTs but instead focuses on enabling capabilities such as enhanced surveillance and electronic flight strip displays. Human factors assessments of these core capabilities have yet to be tested in sufficient detail to determine their benefits potential under stressing environments representative of future air traffic operations. HITLs allow assessment of TFDM's viability under future traffic levels, in controlled environments, under off-nominal conditions and to examine controller acceptance (e.g., workload, situation awareness, trust), while also supporting TFDM assessment across a wide range of objective measures (e.g., performance, environmental, operational error) to help validate and calibrate the results from the computer modeling activities or to fill benefits gaps.

It is proposed to conduct HITLs in the Lincoln Laboratory Airport Simulation Facility (shown in Figure 50) which can be configured to simulate existing tower equipment as well as NextGen-focused "Large Glass" user interfaces such as TFDM. Two system configurations could be examined: (1) a Core TFDM test condition including enhanced surveillance and electronic flight strip systems and (2) a baseline condition including current-day tower systems relevant to the tasks being tested in the HITL (e.g., paper flight strips and existing ASDE-X displays). This direct comparison of TFDM to baseline systems will enable controlled assessments of near-term operational and environmental performance impacts as well as user acceptance of TFDM relative to current operations at any major airport within the NAS. Given the flexibility of the HITL environment, demand (i.e., traffic level) and capacity (i.e., throughput) manipulations can be achieved. Demand, for example, can be manipulated to examine TFDM versus baseline systems under current day and projected 2030 traffic levels.

Figure 50: MIT LL Airport Simulation Facility with Integrated TFDM System

Subtasks would include design and implementation of HITL scenarios, execution of the HITL studies using air traffic control test subjects, and assessment and reporting of the results.

7.2 PROPOSED TASK 2: REFINED DECISION SUPPORT TOOL CAPABILITY ASSESSMENT

Although the "core" TFDM elements that are expected to be the initial focus of TFDM deployment may not include the full set of DSTs, it is strongly recommended that some level of DST computer modeling analysis be maintained even in the short term. This will ensure effective continuation of the interim IID analyses which will be highly relevant for later activities. To perform further computer modeling activities, the three capability areas studied in this report (departure metering, sequence optimization, and airport configuration management) could be extended to explicitly model additional airports and/or model them in more detail. Issues that could be addressed in each of the analysis areas have been highlighted throughout this document. This would provide a higher degree of confidence in benefit projections for the NAS from 2015–2035.

At the direction of the sponsor, additional TFDM DST capabilities could also be examined through computer modeling means. For example, the TFDM decision support areas of departure routing and taxi conformance have not yet been represented in TFDM analysis but could be relevant for future needs. Appropriate models and simulations would be developed and executed to evaluate these additional capabilities. The scope of computer modeling activities can be adjusted flexibly depending on the requirements for data to support TFDM program needs.

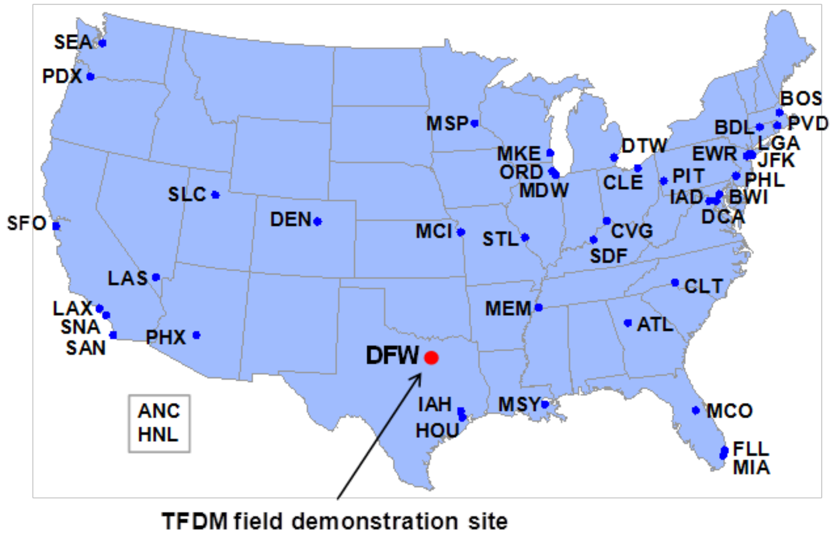
8. SUMMARY

This document has provided an overview of MIT Lincoln Laboratory's activities in support of the interim stage of the Initial Investment Decision benefits assessment for the Tower Flight Data Manager. It has outlined the rationale for the focus areas, and the background, methodology and scope in the focus areas of departure metering, sequence optimization, airport configuration optimization and safety assessment. Estimates of the potential benefits enabled by TFDM deployment has been presented for each of these areas for a subset of airports and conditions considered within the scope of the analyses. These benefits have been monetized where possible. Recommendations for follow-on work, for example, to support the further stages of benefits assessment efforts for TFDM, have also been discussed.

This page intentionally left blank.

APPENDIX A
FULL SET OF TFDM ANALYSIS AIRPORTS

The figure below shows the set of analysis airports assumed to be candidates for "full TFDM" deployment under TFDM-1 (with TIDS, FDM and DSTs). Only a subset of these airports are being considered in the MIT LL computer modeling activities described in this document.



This page intentionally left blank.

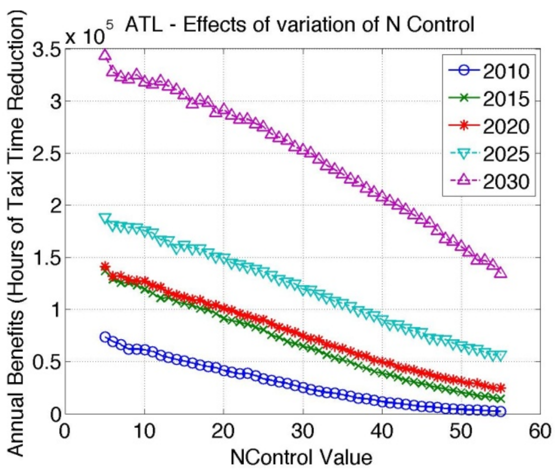
APPENDIX B
DEPARTURE METERING ADDITIONAL INFORMATION

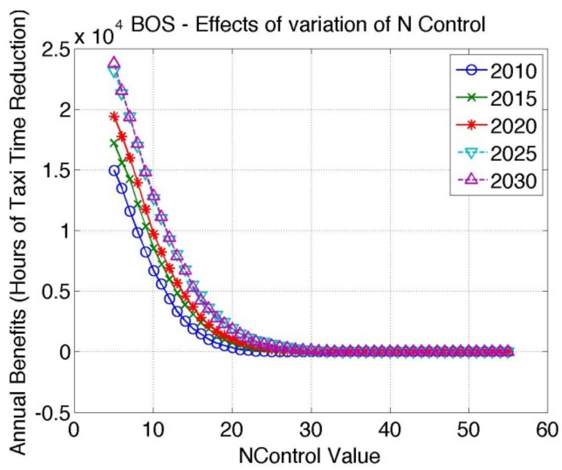
RANDOM FOREST VARIABLES

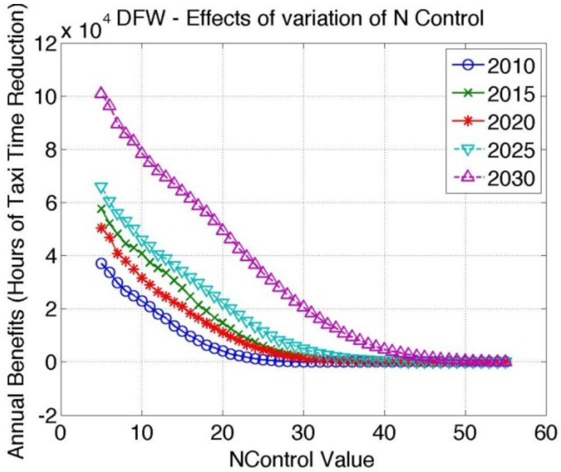
Variable Name	Description	Source
Mean(DepDemand)	Mean(depDemand) - Mean Departure Demand (Yearly, by configuration, per hour)	ASPM - APM
Mean(arrDemand)	Mean(arrDemand) - Mean Arrival Demand (Yearly, by configuration, per hour)	ASPM - APM
90%DepDemand	90%DepDemand - 90th percentile Departure Demand (Yearly, 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 Capacity Used (Yearly, by configuration)	ASPM - APM
# of Arrival Runways	# of Arrival Runways (By Configuration)	Derived from Configuration
# of Departure Runways	# of Departure Runways (By configuration)	Derived from Configuration

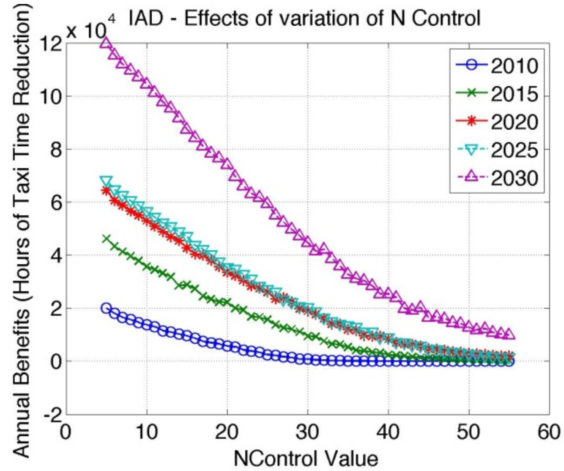
# of Unique Departure Runways	# of Unique Departure Runways (By configuration)	Derived from Configuration
# of Unique Runways	# of Unique Runways (By Configuration)	Derived from Configuration
Area of Airport	Area of Airport (acres)	Wikipedia
Miles of Taxiway	Miles of Taxiway (total)	Wikipedia
Miles of Runway	miles of Runway (total)	Wikipedia
Terminals	Terminals	Wikipedia
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

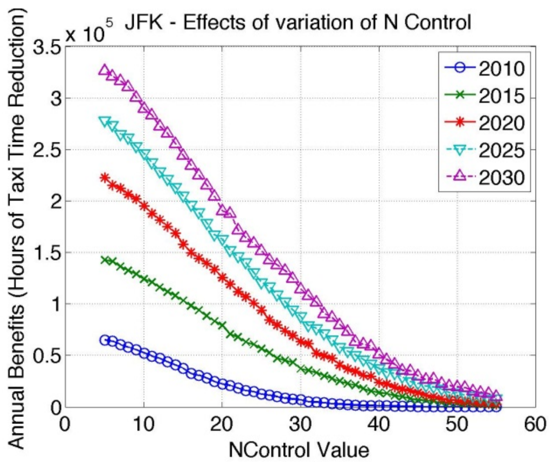
VARIATION OF BENEFITS WITH N CONTROL

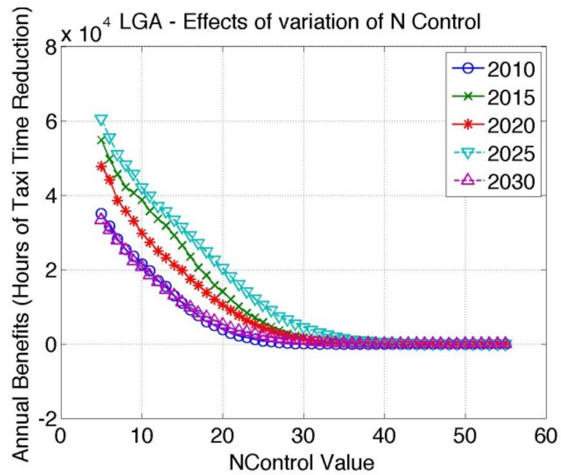


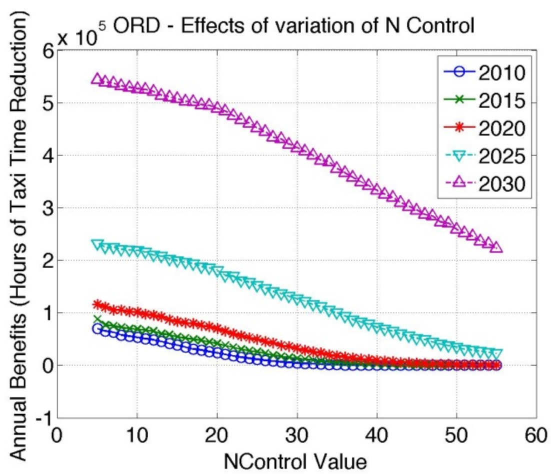


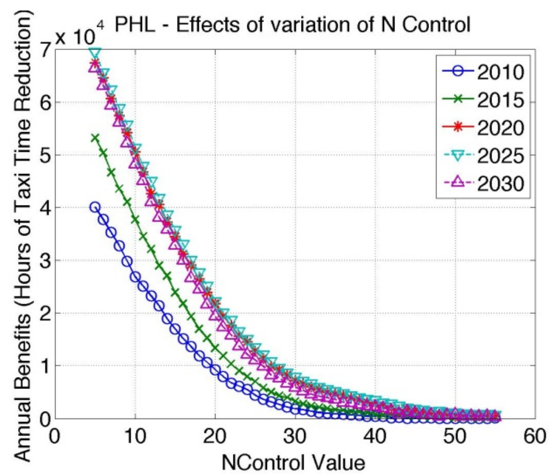

x 10^4 DFW - Effects of variation of N Control
Annual Benefits (Hours of Taxi Time Reduction)
NControl Value
2010
2015
2020
2025
2030


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


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


x 10^4 LGA - Effects of variation of N Control
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


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

This page intentionally left blank.

APPENDIX C
SAFETY ASSESSMENT ADDITIONAL INFORMATION

UNSAFE ACTS

- **Errors:** activities or behaviors that fail to achieve their intended outcome.
    - **Decision errors:** conscious, goal-intended behavior that proceeds as designed, yet proves inadequate or inappropriate for the situation.
    - **Skill-based errors:** errors occurring in the execution of a routine, highly practiced task relating to procedure, training or proficiency that result in an unsafe situation.
    - **Perceptual errors:** errors (such as visual, auditory, proprioceptive, or vestibular illusions, cognitive or attention failures) occurring when sensory input is degraded or unusual.
- **Violations:** activities or behaviors that represent a willful disregard for rules and regulations.
    - **Routine violations:** habitual by nature, often enabled by a system of supervision and management that tolerates such departures from the rules.
    - **Exceptional violations:** isolated departures from authority, neither typical of the individual nor condoned by management.

PRECONDITIONS FOR UNSAFE ACTS

- **Environmental factors:** physical or technological factors that affect practices, conditions and actions of operators and result in human error or an unsafe situation.
    - **Physical environment:** environmental phenomena such as weather or climate conditions that affect the actions of individuals and result in human error or an unsafe situation.
    - **Technological environment:** workspace design factors or automation that affect the actions of individuals and result in human error or an unsafe situation.
- **Conditions of operators:** cognitive, psycho-behavioral, adverse physical states, or physical/mental limitations that affect practices, conditions or actions of individuals and result in human error or an unsafe situation.
    - **Adverse mental state:** mental condition that adversely affect performance (e.g., loss of situational awareness, mental fatigue, pernicious attitudes).
    - **Adverse physiological state:** physiologic event that compromises human performance (e.g., prescribed drugs, physical fatigue, visual illusions).
    - **Physical/mental limitations:** individual lacks the physical or mental capabilities to cope with a situation.
- **Personnel factors:** personal readiness or crew resource management practices, conditions or actions of individuals that result in human error or an unsafe situation.
    - **Crew resource management:** substandard coordination, communication or planning interactions among individuals or teams that result in human error or an unsafe situation.
    - **Personal readiness:** substandard readiness to perform driven by behaviors that disregard readiness regulations or reduce the operating capabilities of the individual.

UNSAFE SUPERVISIONS

- **Inadequate supervision:** failure to identify a hazard, recognize and control risk, provide guidance, training and/or oversight that results in human error or an unsafe situation.
- **Planned inappropriate operations:** failure to adequately assess the hazards associated with an operation thereby allowing for unnecessary risk.
- **Failed to correct problem:** those instances when deficiencies among individuals, equipment, training or other safety areas are "known" to the supervisor, yet are allowed to continue uncorrected.
- **Supervisory operations:** those instances when supervisors willfully disregard existing rules and regulations.

ORGANIZATIONAL INFLUENCES

- **Resource management:** failure in management, allocation, or maintenance of organizational resources, including human resource management, monetary safety budgets, and equipment design that directly or indirectly influences system safety and results in poor error management or creates an unsafe situation.
- **Organizational climate:** situationally-based consistencies (i.e., structure, policies, culture) in the organization's treatment of individuals that influence individual actions and result in human error or an unsafe situation.
- **Organizational process:** formal processes, procedures, and oversight within the organization that result in unrecognized hazards and/or uncontrolled risk and leads to human error or an unsafe situation.

GLOSSARY

AIDS	Accident/Incident Data System
AMASS	Airport Movement Area Safety System
ASDE-X	Airport Surface Detection Equipment, Model X
ASIAS	Aviation Safety Information Analysis and Sharing
ASPM	Aviation System Performance Metrics
ASRS	Aviation Safety Reporting System
ATC	Air Traffic Control
ATL	Hartsfield-Jackson Atlanta International Airport
ATO	Air Traffic Organization
ATO-F	Air Traffic Organization, Finance
BOS	Boston Logan International Airport
CDQM	Collaborative Departure Queue Management
CDS	Collaborative Departure Scheduling
CFR	calls for release
DFW	Dallas/Fort Worth International Airport
DSTs	Decision Support Tools
EDCT	expected departure clearance times
FAA	Federal Aviation Administration
FDM	Flight Data Manager
FID	Final Investment Decision
HFACS	Human Factors Analysis and Classification System
HITL	human-in-the-loop
IAD	Washington Dulles International Airport
IARD	Investment Analysis Readiness Decision
ICAO	International Civil Aviation Organization
IFR	Instrument Flight Rules

IID	Initial Investment Decision
IMC	Instrument Meteorological Conditions
JFK	John F. Kennedy International Airport
LGA	La Guardia Airport
NAS	National Airspace System
NextGen	Next Generation Air Transportation System
NTSB	National Transportation Safety Board
OEDS	Operational Error/Deviation System
OOOI	OUT, OFF, ON, IN
ORD	Chicago O'Hare International Airport
PHL	Philadelphia International Airport
RCCE	runway configuration capacity envelopes
RF	Random Forest
RWSL	Runway Status Lights
SARDA	Spot and Runway Departure Advisory
SOP	Standard Operating Procedures
STBO	Surface Trajectory-Based Operations
SWAC	System Wide Analysis Capability
TAF	Terminal Area Forecast
TFDM	Tower Flight Data Manager
TIDS	Tower Information Display System
TRACON	Terminal Radar Approach Control
VFR	Visual Flight Rules
VMC	Visual Meteorological conditions

REFERENCES

[Balakrishnan & Chandran (2010)] Balakrishnan H. and B. Chandran, "Algorithms for Scheduling Runway Operations under Constrained Position Shifting," *Operations Research*, Vol. 58, No. 6, 2010.

[Barnett et al. (2000)] Barnett, A., G. Paull, and J. Iaedeluca, "Fatal U.S. Runway Collisions Over the Next Two Decades," Air Traffic Control Quarterly, Vol. 8, No. 4, 2000.

[Barnett and Paull (2004)] Barnett, A. and G. Paull, "Effectiveness Analysis for Aviation-Safety Measures in the Absence of Actual Data," Air Traffic Control Quarterly, Vol. 12, No. 3, 2004.

[Bertsimas et al. (2011)] Bertsimas, D., M. Frankovich & A. Odoni, "Optimal Selection of Airport Runway Configurations," *Operations Research,* Vol. 59, No. 6, pp. 1407–1419, 2011.

[Boeing (2011)] Boeing, "Statistical Summary of Commercial Jet Airplane Accidents," retrieved from http://www.boeing.com/news/techissues/pdf/statsum.pdf, 2011.

[Breiman (2001)] Breiman, L., "Random Forests," *Machine Learning*, 45(1):5–32, 2001.

[Brinton et al. (2011)] Brinton, C., C. Provan, S. Lent, T. Prevost & S. Passmore, "Collaborative Departure Queue Management: An Example of Airport Collaborative Decision Making in the United States," *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Berlin, Germany, June 2011.

[Broach & Schroeder (2005)] Broach, D., & Schroeder, D.J., "Relationship of Air Traffic Control Specialist Age to En Route Operational Errors," Final Report DOT/FAA/AM-05/22, Washington, DC, Federal Aviation Administration, 2005.

[Chappell (1994)] Chappell, S.L., "Using Voluntary Incident Reports for Human Factors Evaluation," in *Aviation Psychology in Practice,* N. Johnson, N. McDonald, & R. Fuller (Eds.), Aldershot, England: Ashgate, 1994.

[Damos & Tabachnick (2001)] Damos, D.L., & Tabachnick, B.G., "The Effect of Interruptions on Flight Crew Performance: ASRS Reports," (Technical Report), Los Angeles, CA: Damos Research Associates, Inc., 2001.

[Degani et al. (1991)] Degani, A., Chappell, S.L., & Hayes, M.S. "Who or What Saves the Day? A Comparison of Traditional and Glass Cockpits," *Proceedings of the Sixth International Symposium on Aviation Psychology,* Columbus, OH: The Ohio State University, 1991.

[DoD HFACS (2005)] U.S. Department of Defense, "Department of Defense Human Factors Analysis and Classification System: A Mishap Investigation and Data Analysis Tool," retrieved from http://www.uscg.mil/safety/docs/ergo_hfacs/hfacs.pdf, 2005.

[Endsley (1995)] Endsley, M. R. (1995), "A Taxonomy of Situation Awareness Errors", in *Human factors in Aviation Operations: Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP; pp. 287–292)*, Gower Technical, 1995.

[FAA (2010)] Federal Aviation Administration, "Shortfall Analysis Report for the Tower Flight Data Manager (TFDM) Program," July 2010.

[FAA/APO (2011)] FAA Office of Policy and Plans (APO), "Terminal Area Forecast (TAF) Data," retrieved from http://aspm.faa.gov/main/taf.asp, 2011.

[FAA/ATO (2011)] FAA, "Economic Information for Investment Analysis Data Package," March 2011.

[FAA/ASPM (2011)] FAA, "Aviation System Performance Metrics," http://aspm.faa.gov, 2011.

[FAA/MITRE (2007)] 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.

[FAA/TAF (2011)] FAA, "Terminal Area Forecast Summary: Fiscal Years 2010–2030," http://aspm.faa.gov/main/taf.asp, 2011.

[FAA/Weather (2010)] FAA, "Weather-related Aviation Accident Study 2003–2007," Study Report, Washington, DC: Federal Aviation Administration, 2010.

[Funk et al. (1999)] Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & Owen, G., "Flight Deck Automation Issues," *International Journal of Aviation Psychology*, 9(2), 109–123, 1999.

[Gilbo (1993)] Gilbo, E.P., "Airport Capacity: Representation, Estimation, Optimization," *IEEE Transactions on Control Systems Technology*, Vol. 1, No. 3, pp. 144–154, 1993.

[Gilbo (1997)] Gilbo, E.P., "Airport Capacity Utilization," *IEEE Transactions on Control Systems Technology*, Vol. 5, No. 5, pp. 490–503, 1997.

[Gupta et al. (2009)] Gupta, G., W. Malik & Y.C. Jung, "A Mixed Integer Linear Program for Airport Departure Scheduling," *AIAA Aviation, Technology, Integration, and Operations (ATIO) Conference* and *Aircraft Noise and Emissions Reduction Symposium (ANERS)*, Hilton Head, SC, AIAA-2009-6933, 2009.

[Gupta et al. (2010)] Gupta, G., W. Malik & Y. C. Jung, "Incorporating Active Runway Crossings in Airport Departure Scheduling," *AIAA Guidance, Navigation, and Control (GNC) Conference* and *Modeling and Simulation Technologies (MST) Conference*, Toronto, Canada, 2010.

[Hoang et al. (2011)] Hoang, T., Y.C. Jung, J. Holbrook & W.A. Malik, "Tower Controllers' Assessment of the Spot and Runway Departure Advisor (SARDA) Concept," *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Berlin, Germany, June 2011.

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

[Johnson (2005)] Johnson, B, "Business Case Analysis Report for Airport Surface Detection Equipment, Model X (ASDE-X)," Exhibit 300, Attachment 2, 2005.

[Jones & Endsley (1995)] Jones, D.G., & Endsley, M.R., "Investigation of Situation Awareness Errors," *Proceedings of the Eighth International Symposium on Aviation Psychology*, Columbus, OH: The Ohio State University, 1995.

[Kang et al. (2007)] Kang, L., M. Berge and M. Carter, "Selecting Representative Weather Days for Air Traffic Analysis," The Airline Group of the International Federation of Operational Research Societies (AGIFORS) Annual Symposium, 2007.

[Lee et al. (2010)] Lee, H., I. Simaiakis & H. Balakrishnan, "A Comparison of Aircraft Trajectory-based and Aggregate Queue-based Control of Airport Taxi Processes," *IEEE Digital Avionics Systems Conference*, 2010.

[MIT LL (2010)] MIT Lincoln Laboratory, "Field Demonstration #1 Final Report for Tower Flight Data Manager (TFDM) and Staffed NextGen Tower (SNT)," Nov. 15, 2010.

[MIT LL (2011)] MIT Lincoln Laboratory, "Field Demonstration #2 Final Report for Tower Flight Data Manager (TFDM)," Dec. 12, 2011.

[MITRE (2010)] Audenaerd, L.F., C.S. Burr, P.A. Diffenderfer & C.E. Morgan, "Surface Trajectory-Based Operations (STBO) Mid-Term Concept of Operations Overview and Scenarios," Revision 1, MITRE Document MP090230R1, 2010.

[Nakahara *et al.* (2011)] 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.

[Noonan (2011)] Noonan, K., "FAA's System-Wide Analysis Capability," Integrated Communications, Navigation and Surveillance Conference (ICNS), 2011, 10–12 May 2011, doi:10.1109/ICNSURV.2011.5935365, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5935365&isnumber=5935117

[Nowinski et al. (2003)] Nowinski, J.L., Holbrook, J.B., & Dismukes, R.K., "Human Memory and Cockpit Operations: An ASRS Study," *Proceedings of the Seventh International Symposium on Aviation Psychology*, Columbus, OH: The Ohio State University, 2003.

[Reason (1990)] Reason, J., *Human Error*, New York: Cambridge University Press, 1990.

[Rodgers (1993)] Rodgers, M.D., "An Examination of the Operational Error Database for Air Route Traffic Control Centers," Final Report DOT/FAA/AM-93/22, Washington, DC: Federal Aviation Administration, 1993.

[Simaiakis (2009)] 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.

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

[Simaiakis et al. (2011)] Simaiakis, I., H. Khadilkar, H. Balakrishnan, T.G. Reynolds, R.J. Hansman, B. Reilly & S. Urlass, "Demonstration of Reduced Airport Congestion through Pushback Rate Control," *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*, Berlin, Germany, 2011.

[Solveling et al. (2010)] Solveling, G., J-P. Clarke, E. Johnson & S. Solak, "Runway Operations Optimization in the Presence of Uncertainties," *10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Fort Worth, Texas, AIAA-2010-9252, 2010.

[Wiegmann & Shappell (2003)] Wiegmann, D.A., & Shappell, S.A., "A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System," Burlington, VT: Ashgate Publishing Company, 2003.

[Wiegmann and Shappell (2001)] Wiegmann, D.A. and S.A. Shappell, "Human Error Analysis of Commercial Aviation Accidents: Application of the Human Factors Analysis and Classification System (HFACS)," Aviation, Space, and Environmental Medicine, Vol. 72, No. 11, 2001.

[Wilhelmsen (1994)] Wilhelmsen, H., "Preventing Runway Conflicts: The Role of Airport Surveillance, Tower-cab Alerts, and Runway Status Lights," *The Lincoln Laboratory Journal*, 7(2), 149–168, 1994.