

# Benefits Assessment of Integrating Arrival, Departure, and Surface Operations with ATD-2

Aditya Saraf, Valerie Sui, Kennis Chan, Natasha Luch, Martin Popish, Evan Lohn, Brandon Huang  
ATAC Corporation  
Santa Clara, CA, USA  
aps@atac.com

Benjamin Levy and Marc Rose  
MCR Federal Inc.  
Washington, D.C., USA  
blevy@mcri.com

Hamsa Balakrishnan  
Massachusetts Institute of Technology  
Cambridge, MA, USA  
hamsa@mit.edu

Husni Idris  
TASC/Engility  
Billerica, MA, USA  
husni.idris@nasa.gov

**Abstract**— The combined transit of departure flights from the airport surface, through the terminal airspace and merging into overhead en route traffic streams is a major source of delay in the National Airspace System (NAS). This is especially true in metroplex regions where departures and arrivals from/to multiple, proximate airports compete for limited resources (e.g., mixed-use runways, shared departure-fixes, busy overhead traffic streams). Current-day metroplex traffic management practices lead to multiple shortfalls. Under the Airspace Technology Demonstration – 2 (ATD-2) subproject, NASA plans to address these shortfalls by demonstrating Integrated Arrival, Departure, Surface (IADS) technologies and transitioning them to field-implementation. These technologies aim to increase the predictability, efficiency, and throughput of metroplex operations while meeting future air traffic demand. In order to help with effective transition of ATD-2 tools to the field, NASA needs reliable information regarding the operational shortfalls that ATD-2 can address, its benefit mechanisms and relevant performance metrics as well as high-fidelity benefit-cost estimates of implementing the ATD-2 system at major airports in the NAS. This paper describes the results of modeling, simulation and data analysis work performed in order to develop reliable estimates of the benefits of NASA’s ATD-2 concept on a nationwide scale based on high-fidelity, realistic models of ATD-2 performance.

**Keywords**—*integrated arrival departure surface operations, time-based scheduling, departure metering, air traffic management*

## I. INTRODUCTION

The management of departure flights taking off from the airport surface, traveling through the terminal airspace and merging into overhead en route traffic streams presents a complex scheduling and air traffic management (ATM) problem. This is especially true in metroplex regions where departures and arrivals from/to multiple, proximate airports compete for limited resources (e.g., mixed-use runways, shared departure-fixes, busy overhead traffic streams). Current-day metroplex traffic management practices lead to multiple operational shortfalls. These shortfalls include: (i) Identical ticketed departure times, a pushback-when-ready operational paradigm, and reactive first-come-first-served (FCFS) management of clearances at ramp transition spots, lead to inefficient departure sequences causing taxi inefficiency (stop-and-go) and throughput loss; (ii) Lack of predictability in the departure process forces tower controllers to impose buffers

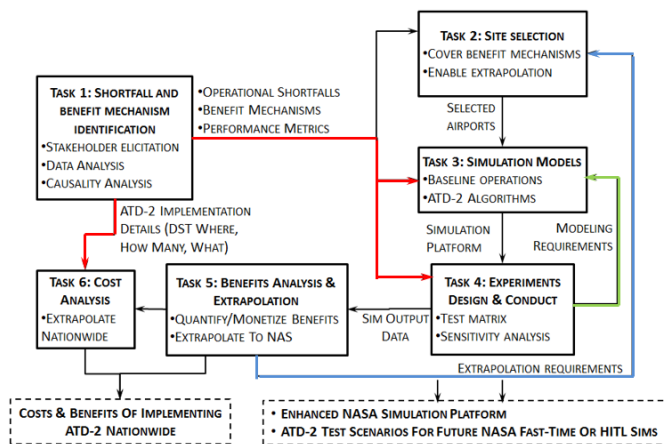
(e.g., runway separation buffers) to ensure safety; and forces the receiving Terminal Radar Approach Control (TRACON) facility and Centers to impose inefficient departure restrictions, (e.g., excess miles-in-trail (MIT), or approval requests, (APREQs)) on airports to make space for airborne merging; and (iii) Lack of predictability also causes airlines to set excess block times, which limits fleet utilization and increases personnel and fuel costs.

Under the ATD-2 subproject, NASA plans to address these shortfalls by demonstrating Integrated Arrival, Departure, Surface (IADS) technologies that comprise the ATD-2 system and transitioning them to field-implementation. These technologies aim to increase the predictability, efficiency, and throughput of metroplex operations while meeting future air traffic demand [1]. The operational environment for the ATD-2 system consists of a local metroplex airspace overlying one or more well-equipped airports (e.g., busy airports with surface surveillance radars installed) and multiple less-equipped airports (e.g., smaller, less busy airports without surface surveillance systems). Departures from these airports may share departure fixes on the TRACON boundary and merge into busy en route traffic streams in the Center airspace. Departures are subject to multiple restrictions including MITs at en route merge points and departure fixes; Expected Departure Clearance Times (EDCTs) from Ground Delay Programs; Weather-related departure fix/gate closures; and Takeoff time restrictions due to arrival metering constraints at a destination airport. The ATD-2 system will compute time-based departure schedules for all airports in the local metroplex while accounting for all departure restrictions.

In order to help with the transitioning of ATD-2 tools to the field, NASA needs reliable information regarding the operational shortfalls that ATD-2 can address, its benefit mechanisms and relevant performance metrics as well as high-fidelity benefit-cost estimates of implementing the ATD-2 system at NAS-wide airports. The overarching objective of the research work described in this paper is to fulfil this need by generating high-fidelity benefit and cost estimates of implementing NASA’s ATD-2 system at major airports in the NAS. This paper provides results from initial efforts to develop reliable estimates for the benefits part of the equation.

We take a simulation-based approach for estimating ATD-2 benefits. To support this approach, we have developed a high-fidelity simulation environment for simulating aircraft trajectories in both the surface and airspace subsystems of the ATD-2 system, under current-day ATM procedures as well as under ATD-2 procedures, and then compared the performance metrics under these two procedures. We take a mixed fidelity modeling approach where different domains within the ATD-2 subsystems are modeled at different levels of fidelity depending upon site-specific operational characteristics and the specific benefit mechanism being analyzed. Our simulation models simulate all the key constraints faced by departure flights along their path from gate to en route traffic stream merging. These constraints include traffic congestion in the ramp area and on movement area taxiways at the departure airport, capacity constraints on the departure airport runway system, capacity constraints at departure-fixes where flights from multiple TRACON airports merge, and constraints at entry points to overhead en route traffic streams.

Figure 1 shows our overall technical approach. Our approach was to conduct high-fidelity simulation-based benefits assessments at three selected airport sites and then extrapolate the results to NAS-wide and annualized benefits estimates. Task 3 was the core simulation environment development task. Our aim is to be comprehensive in modeling all current-day operational shortfalls and the alleviating impact of ATD-2 on them. Hence, before generating the high-fidelity simulation environment, in Task 1, we performed a comprehensive identification of (i) operational shortfalls that ATD-2 can address, (ii) associated ATD-2 benefit mechanisms, and (iii) metrics that can be used to measure the benefits of ATD-2. Then, in Task 2 we performed a historical operations data analysis covering 35 top U.S. airports and used subject matter expert inputs for selecting three airport sites as candidates for detailed simulation modeling.



**Figure 1. Overall Technical Approach for Assessment of NAS-wide ATD-2 Benefits and Costs**

Based on Task 2 analysis, we selected Newark Liberty International (EWR), Dallas Fort Worth International (DFW), and Charlotte/Douglas International (CLT) as the three airports to be studied. The selected airports cover all the important

ATD-2 related operational features identified in Task 1 as outlined below:

- EWR has a unique runway interaction geometry with one pair of closely spaced parallel runways as well as an intersecting runway, with taxiing arrivals required to cross an active departure runway to reach the terminal gates. EWR displays high variability in taxi-out times and gate pushback times as well as a high level of taxi-out delay and taxi-stop duration. Our analysis of airport runway capacity saturation showed that EWR operates under departure capacity saturation for around 80% of the time. Only two airports (LGA and PHL) out of the 35 airports in our analysis set spent higher percentages of time under saturation conditions than EWR. Also, EWR experiences higher local queuing delays relative to delay generated due to downstream restrictions. EWR was also the 2nd ranked airport in terms of the estimated potential to benefit from departure metering technologies as per our analysis of taxi-out delay and fuel savings [2]. Moreover, EWR displayed high levels of departure-fix sharing, departure flight altitude level-off inefficiencies, and local departure restrictions severity/frequency.
- DFW differs significantly from EWR in terms of its airport geometry and departure airspace configuration. Its runway system has relatively lower interaction-impacts: the only dependencies are between arrivals and departures using closely-spaced parallel runways, and arrivals crossing an active departure runway to reach their gates. DFW displays medium levels of taxi-out delay, taxi stop duration, taxi-time variability, gate pushback time variability, and time spent under departure capacity saturation. DFW displayed a very low taxi-out time savings potential and medium fuel savings potential to benefit from departure metering technologies. External departure restrictions and departure flight level-offs were identified as significant constraining factors in the D10 TRACON airspace. D10 displayed only a medium level of departure-fix sharing between DFW and other neighboring airports.
- CLT displays an interesting but different runway interaction geometry—with independent parallel runways, an intersecting arrival runway, a mixed-use departure-heavy runway and a large ramp area. In our analyses, CLT displayed high taxi-out time variability, low gate pushback time variability, high taxi-out delay, high taxi stop duration, a medium percentage of time spent under departure capacity saturation, and a medium level of potential to benefit from departure metering in terms of taxi-out time savings and fuel savings. In terms of airspace constraints, CLT has almost non-existent departure-fix sharing, a small degree of TRACON departure altitude level-off inefficiencies, but a high level of external tactical departure scheduling constraints coming from neighboring Centers as well as from Time-Based Flow Management (TBFM) metering for the busy ATL arrival traffic stream.

After selecting three simulation candidate airports in Task 2, we developed the high-fidelity simulation environment for these airports in Task 3. Task 4, which we have not finished so far, involves conducting high-fidelity simulation based assessments at these three airport sites using a manageable number of traffic scenarios per site, and analyzing the simulation outputs to compute benefit metrics. Task 5, which is another future task, then extrapolates the benefits results to a wider set of airports (FAA Core 30 airports) and to an annualized scale. Task 6, another future task, involves analyzing the cost of implementing ATD-2 system at nationwide airports.

The rest of this paper is structured as follows. Section II describes the hybrid, surface-airspace simulation environment that we developed to support our benefits assessments. Section III describes simulation experiment design steps that we performed. Section IV presents results from the initial simulations and discusses the main findings. Section V presents our approach for extrapolating these high-fidelity simulation results to nationwide and annualized benefits estimates. Finally, Section VI summarizes the main conclusions from research work presented in this paper.

## II. SIMULATION ENVIRONMENT

In order to support simulation-based benefits assessments of the ATD-2 system we generated a high-fidelity simulation environment. The core of our simulation environment is NASA’s high-fidelity Surface Operations Simulator and Scheduler (SOSS) simulation platform [3-5]. SOSS simulates departure and arrival flight trajectories on the airport surface. We use SOSS to simulate traffic on the surface of the selected ATD-2 airport sites. We integrate SOSS with the ATAC Airspace Operations Simulator and Scheduler (AOSS), which is a MATLAB-based airspace traffic queuing simulation. AOSS complements SOSS by simulating aircraft trajectories in the TRACON and en route airspace along a node-link network of frequently-flown airspace routes.

NASA’s SOSS is a fast-time simulation platform used to simulate airport surface operations and support rapid prototyping of surface scheduling algorithms. SOSS includes a high-fidelity node-link model of airport gates, taxi paths, and runways on the airport’s surface. It also includes trajectory-based models for simulating aircraft moving on the airport surface, with aircraft type-specific surface transit speed modeling and special runway speed transit modeling. SOSS also includes models of pilot self-separation which prevents flights from getting too close to each other. SOSS was not designed to be a standalone modeling tool. SOSS was designed to be used in conjunction with external scheduling components. When integrated with external schedulers it is SOSS’s job to move aircraft on the surface according to the recommended schedule, and monitor separation violations and scheduling conformance. In our simulation environment, AOSS acts as an external scheduler to SOSS.

Figure 2 shows the interconnected SOSS-AOSS system. As shown in the figure, SOSS transfers over the simulation-control

of a departure flight to AOSS when the departure flight takes off. AOSS then simulates the movement of the departure flight along its airborne route from takeoff runway to departure fix and then on to an en route stream merge point. AOSS includes queuing simulation-based models of the departure-fix merge process as well as the en route stream merge process. In addition to the focus ATD-2 airport departures, AOSS also simulates departures from satellite airports within the same TRACON as well as departures from NAS-wide airports that merge with the focus ATD-2 airport’s departures in the en route airspace.

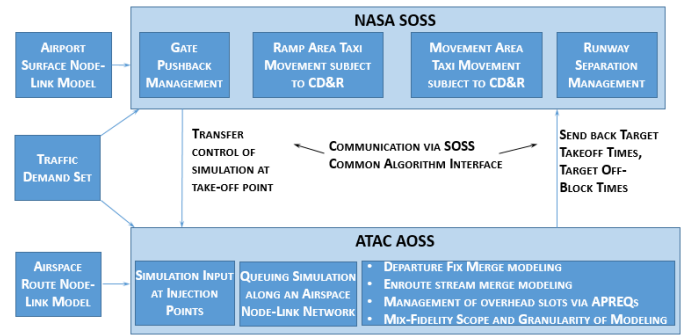
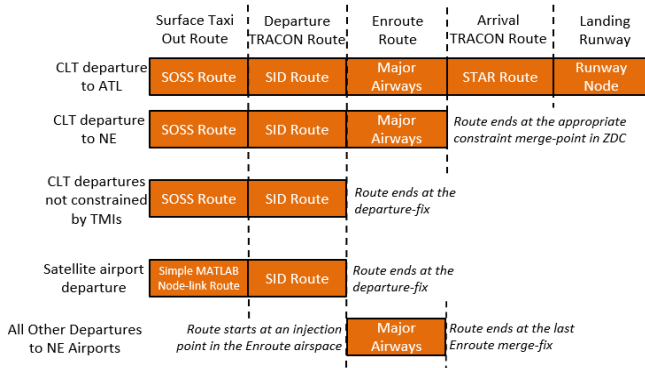


Figure 2. Integrated Surface Airspace Simulation Environment

AOSS takes a mixed fidelity modeling approach, whereby different domains within the airspace transit of departure flights are modeled at different levels of fidelity depending upon site-specific operational characteristics and the specific benefit mechanism being modeled/analyzed. Figure 3 shows the different levels of modeling fidelity used in AOSS for simulating different operation types. As shown in the figure, CLT departure to ATL are treated as special. Their trajectories are modeled by nodes and links starting at the departure runway and ending at the landing runway at ATL. Trajectories for CLT departures going to the U.S. Northeast Corridor airports, which are most frequently impacted by traffic management initiatives (TMIs), are modeled starting at the departure runway and continuing on to their en route merge point, usually falling within Washington D.C. Center (ZDC) airspace. For all other CLT departures, the modeled airspace routes end at the departure-fix merge. We also model routes for departures from CLT satellite airports. For these departures, the routes are modeled from their departure runway until the departure-fix merge. We also model all NAS-wide departures (i.e., non-CLT and non-CLT satellite departures) that interact with CLT departures going to the Northeast Corridor airports at some point in their transit. Modeling these “other” departures serves to create realistic en route slot fitting constraints for CLT departures and improves the simulation fidelity. These “other” departures enter the simulation at user-defined injection points which usually fall on Center boundaries. Their airspace routes end at the respective en route merge fixes.

AOSS simulation components are designed to evaluate the operational impacts and benefit mechanisms of ATD-2 decision support tools in comparison with current-day departure operations under varying levels of departure

restrictions. The key AOSS components are (i) An Airspace Node-link Network: The AOSS airspace node-link network consists of the airspace routes most commonly used by departures from the main ATD-2 focus airport and its satellite airports with which the main airport departures share TRACON boundary departure fixes, as well as all other NAS-wide traffic that interacts with the main airport departures in en route airspace., (ii) Transit Time Models: AOSS includes transit time models for all the links within the node-link network. These transit times are derived from analysis of historical operations data, and (iii) Queue Control at Nodes: AOSS also contains a queue control module for simulating the merging of aircraft at key nodes over the node-link network. This queue control module provides a first-cut estimate of airborne delays introduced by sequencing for merging at the TRACON departure fixes and for merging into the overhead en route traffic streams.



**Figure 3. Different Scopes and Fidelities of Departure Airspace Trajectory Modeling**

### III. SIMULATION SCENARIO DESIGN

In our work we paid careful attention to designing appropriate simulation scenarios in order to achieve the inclusion of a broad variety of key influencing factors in our simulation experiments. Carefully chosen scenarios enable us to reliably extrapolate benefits results from a small number of representative simulation experiments to nationwide and annualized benefits estimates. The most important factor in scenario design is selecting appropriate historical dates for basing the simulation scenarios on. Weather and demand are the principal factors which govern simulation date selection. Our simulation day selection methodology determines six conditions under which ATD-2 is expected to provide differing levels of benefits. We determine these conditions by looking at key weather and traffic demand factors that measure the performance of departure metering: departure demand, departure delay, local and national weather impacts, and occurrence/severity of TMIs. One historical day fitting each condition is selected for simulation scenario development. Table 1 shows the conditions and expected outcome of benefits estimation from the model, as well as the selected dates that match with each condition.

Conditions on 5/6/16 provided an expected upper bound of benefits achievable from the ATD-2 model. Demand is high and TMI constraints are imposed to manage demand-capacity imbalance, but weather conditions are moderate near KCLT airport and across the CONUS. As such, these conditions provide an upper bound on the benefits, scaled by frequency of occurrence and other modeling efforts.

**Table 1. Simulation Date Selection Conditions**

DATE	DEP DEMAND AT KCLT	DAILY DELAY AT KCLT		WEATHER IMPACT		TMI CONSTRAINTS		EXPECTED ATD-2 BENEFIT
	AT KCLT	DEPART URE	TAXI -OUT	CONUS	CLT	APREQ INDEX	MIT INDEX	
1/9/16	Low	Mod	Low	Low	Low	Low	N/A	Low
5/6/16	High	Mod	High	Mod	Mod	Mod	Mod	Highest
6/17/2016	High	High	Mod	High	High	Mod	Mod	Moderate
8/30/2016	Low	Low	Mod	High	High	High	N/A	Low
6/28/2016	Low	Mod	High	Mod	Mod	Low	High	Moderate
8/1/2016	Mod	High	High	High	Mod	High	High	Moderate

On 6/17/16, demand is high and constraints are imposed to manage demand-capacity imbalance, where capacity is reduced by poor weather conditions. For these conditions, we expect that the benefits of the ATD-2 model to be moderate, since the implementation of the ATD-2 technology under these conditions can only mitigate so much delay; in other words, the conditions are poor, but some delay is recoverable.

Low benefits on 1/9/2016 should be expected, given ‘blue-sky’ conditions, low demand, and few constraints. Similarly, with low demand under poor weather conditions (i.e., 8/30/16), low model benefits should result; this condition serves as an additional check on the model validation, as the model should make no unnecessary recommendations or changes.

Finally, we propose two simulation dates on which the APREQ and MIT impact indices were high, respectively, as 6/28/16 and 8/1/16.

Two independent date selection methodologies were developed and integrated to provide the simulation dates, for the conditions described in Table 1. Under the first approach, we present the methodology for assessing the combined impact of weather and demand on as-flown flights for 20 canonical dates for fiscal year 2016 (10/12/2015 to 9/23/2016), as developed by the US Federal Aviation Administration [6], and the two additional dates of high APREQ and MIT impact indices. The second approach is based on analysis of TMI constraints in National Traffic Management Log (NTML) data.

#### A. Weather Impact Traffic Index Approach

Based on the weather-impact traffic index (i.e., WITI) construct developed by [7], we extended and applied the original methodology to estimate the WITI score for 22 representative dates in the selection pool. We adapted the original WITI methodology from a CONUS-wide assessment to allow application of smaller-scale regions (e.g., Centers, regions around the simulation airports of CLT, DFW, and EWR).



**B. APREQ and Miles-in-Trail Impact Methodology**

An analysis of historical departure restrictions was conducted to support the simulation date selection analysis by providing data on the extent and severity of departure restrictions occurring each day during FAA fiscal year 2016. Our analysis looked at two types of departure restrictions—MITs and APREQs—imposed on departures taking off from the focus model airports (CLT, EWR and DFW). We analyzed one year’s worth of NTML data for the FAA fiscal year 2016 to support this analysis. Our aim was to select days with different levels of extent of MIT and APREQ occurrence (e.g., number, duration and size of restrictions) as well as different levels of the severity of their impact (e.g., departure delay impact) on the focus airport departure flows. To support this aim we developed two measures or indices to help us with our analysis—(i) a MIT Restriction Impact Index, which is a measure of how severely an airport was impacted by MIT restrictions on a particular day, and (ii) an APREQ Restriction Impact Index, which is a measure of how severely an airport was impacted by APREQ restrictions on a particular day. Both of these indices are computed for each focus airport per day by analyzing NTML data (to compute the time-extent of the restriction) in conjunction with end-to-end track data (to identify which flights are impacted by the restriction and to compute their respective departure delays). The methodology for computing these indices is found in [8]. Table 2 presents the APREQ and MIT index scores for the analysis dates considered in this study.

**Table 2. APREQ and MIT Index Scores for CLT Airport**

DATE	DEPARTURES FROM KCLT	APREQ IMPACT INDEX	MIT IMPACT INDEX
10/12/2015	740	2090	40
10/13/2015	730	705	165
11/11/2015	747	808	N/A
12/4/2015	734	N/A	N/A
12/27/2015	735	200	0
01/09/2016	656	28	N/A
1/24/2016	570	23	0
3/3/2016	744	N/A	790
3/6/2016	712	483	165
3/17/2016	754	325	N/A
4/7/2016	761	545	344
5/6/2016	763	478	N/A
5/21/2016	685	553	234
6/17/2016	756	185	775
6/27/2016	736	473	270
7/18/2016	757	448	1748
8/4/2016	751	2055	N/A
8/30/2016	649	5215	N/A
9/10/2016	590	1723	N/A
6/28/2016	733	468	3507
8/1/2016	745	6053	2495
9/23/2016	736	3150	N/A

**C. Methodology Integration and Simulation Date Selection**

Separately, two methods of assessing constraints on departures from CLT Airport exist, as described in Section III.A Section III.B. The integration methodology presented below provides the selection of simulation dates according to the format and criteria listed in Table 1. As such, we create a

cumulative mass distribution and its rank for the following relevant criteria per day at each focus airport:

- Number of ASPM departures
- APREQ Impact Index
- MIT Impact Index
- Total daily departure delay (minutes)
- Total daily taxi-out delay (minutes)
- Daily WITI scores for all geographic regions

The use of the rank statistics for the above-listed metrics allows us to choose the dates which satisfy the differing conditions of demand, constraint (e.g., APREQ, MIT), and weather. In particular, the geographical WITI scores allow the identification of dates with low weather impacts and high demand, both at the CONUS-scale and in the locality of CLT Airport. The WITI scores per region also allow the separation of local traffic impact on departures from CLT and the extrapolation of the impact of CONUS-wide weather and demand on the departure and taxi-out delays at CLT Airport.

For each of the dates in the selection pool, we created a cumulative mass likelihood (i.e., sample probability of exceedance). The cumulative mass likelihood for each metric-date combination was ranked, with rank = 1 for the highest cumulative mass likelihood value per metric and rank = 22 for the lowest (equal to the number of days in the selection pool). Finally, we devised categorical variables for each metric based on rank ranges, where  $1 \leq \text{rank} \leq 7$  demarked ‘High’,  $8 \leq \text{rank} \leq 14$ , denoted ‘Moderate’, and  $15 \leq \text{rank} \leq 22$  indicated ‘Low’. On 5/6/2016, the conditions were high departure demand at KCLT, moderate to high daily delay at CLT (departure delays and taxi-out delays), moderate weather impact across the CONUS and over CLT, and moderate TMI constraints (based on APREQ and MIT impact indices). The results of the ranking and categorical classifications are presented in Table 3 below.

**Table 3. Ranking and Categorical Variable Results for Simulation Dates, (Color Key: Base Selected Dates – faint yellow background, Alternate Dates – dark background)**

DATE	TMI IMPACT INDEX		WEATHER IMPACT		DEPARTURE	DAILY TOTAL DELAYS	
	APREQ	MIT	CONUS	CLT	DEMAND	DEPARTURE	TAXI-OUT
10/12/2015	High	Mod	Mod	High	Mod	Low	Low
10/13/2015	Mod	Mod	Mod	High	Low	Low	Low
11/11/2015	High		Mod	High	High	Low	Mod
12/4/2015			Low	Low	Mod	Low	Mod
12/27/2015	Low	Mod	Low	Low	Mod	High	Low
1/9/16	Low		Low	Low	Low	Mod	Low
1/24/2016	Low	High	Low	Low	Low	High	Low
3/3/2016	Mod	Mod	Low	Low	Mod	Mod	Low
3/6/2016	Low		Low	Low	Low	Low	Low
3/17/2016	Mod	High	Mod	Mod	High	Mod	Mod
4/7/2016	Mod		Mod	Mod	High	Mod	High
5/6/16	Mod	Mod	Mod	Mod	High	Mod	High
5/21/2016	Low	High	Low	Low	Low	High	Mod
6/17/2016	Mod	Mod	High	High	High	High	Mod
6/27/2016	Low	High	High	Mod	Mod	High	High
7/18/2016	Mod	High	High	Mod	High	High	High
8/4/2016	High	Mod	High	High	High	Mod	High
8/30/2016	High		High	High	Low	Low	Mod
9/10/2016	High		Low	Low	Low	Low	Low
6/28/2016	Low	High	Mod	Mod	Low	Mod	High
8/1/2016	High	High	High	Mod	Mod	High	High
9/23/2016	High		High	High	Mod	Low	Mod

In summary, the ranking based analysis of prevailing traffic demand, TMI and weather-related constraints on each selection pool day led us to select the simulation days identified in Table 1 (i.e., the same as the highlighted dates in Table 3).

#### IV. SIMULATION EXPERIMENT VALIDATION AND RESULTS

In this section we present results from an initial set of simulations that we performed using the simulation environment discussed in the previous Section. The purpose of these initial simulations was to validate that our simulation platform is accurately simulating current-day (baseline) operations as well as operations controlled by the ATD-2 scheduler. Special focus was on verifying that the simulation platform is correctly simulating the impacts of en route APREQ constraints on the airport surface traffic.

##### A. Initial Simulation Experiment Environment

The focus of our initial set of simulation experiments was on CLT airport, using a simulation scenario derived from real, historical operations data from a busy, high-delay day of traffic. We selected 7/21/2016 as our candidate simulation day for these initial simulations. This was a high traffic, high departure delay day. We started our initial simulation experiments before we had the finalized list of selected dates (see Section III.C), hence we selected an initial simulation day outside that list. We plan to conduct more simulations using the selected days listed in Table 1.

In the initial simulations presented here, we modeled a South-flow configuration at CLT airport. We simulated traffic from 0900 UTC to 1700 UTC on 07/21/2016, which included three departure pushes and two arrival pushes. The traffic scenario for these simulations was derived from ASDE-X surveillance data, which was available to us via NASA's Sherlock ATM data warehouse [9]. Surveillance data was augmented with gate-allocation and actual gate-out time information. The Sherlock ATM data warehouse is a crucial piece of the ATM research infrastructure used by NASA Ames and its partners. Sherlock comprises several components, including a database of raw data collected from various NAS systems, parsed and processed data, derived data, and reports derived from pre-defined queries, as well as a Web-based user interface, and supplementary services for query and visualization.

The CLT South-flow model that we used for these simulations has undergone extensive validation both by both NASA simulation experts [10] as well as by our research team [8], and we are confident that our surface model is simulating airport surface operations correctly, except for a minor discrepancy in predicting movement area taxi-out times [8].

In our simulations, we also modeled APREQ constraints on CLT departures going to Northeast Corridor airports using two busy en route merge fixes in the ZDC airspace – BEGVE and MALNR. Many CLT departures going to Northeast Corridor destination airports (e.g., LGA, EWR, PHL, TEB) merge into overhead en route traffic streams at or near these two fixes. In

our simulations, these CLT departure flights received APREQ departure time constraints which they had to adhere to (i.e., they had to takeoff within a  $-3/+2$  minute window around a target takeoff time for hitting a gap at the en route merge fix). The APREQ constraint was active from 1000 UTC to 1700 UTC in the simulations.

We simulated both current-day operations (i.e., baseline) as well as operations under the control of the ATD-2 scheduler, in our simulations. Under simulated current-day operations some of the APREQ-impacted flights received takeoff time window constraints while they were at their gates, which caused them to hold at the gates longer to hit the APREQ windows. Other APREQ-impacted flights received the constraints after they had pushed back, in which case they had to absorb the required delay airing taxi or miss their APREQ window. Under simulated ATD-2 operations, all departure flights, including the APREQ-impacted flights, received target off block times (TOBTs) so as to delay them at the gate far enough to meet APREQ target takeoff time constraints as well as runway system capacity constraints and departure-fix merge capacity constraints. These TOBTs were computed by an in-house version of NASA's ATD-2 scheduler [11] that we developed. We used a maximum gate delay limit of 5 minutes for all flights that were held back at their gates. Note that these were our very first round of simulations and the parameter settings (gate delay limit, etc.) discussed above were selected based on limited knowledge of actual operational factors. We plan to conduct interviews with subject matter experts in the upcoming months to learn more about the current-day handling of APREQs as well as the future APREQ handling procedure under ATD-2, and will update the simulation settings accordingly. Moreover, the capabilities of our ATD-2 scheduler also do not match exactly with NASA's version of the ATD-2 scheduler because both the schedulers were developed independently. We plan to make changes to our scheduler logic to bring it closer in functionality to NASA's scheduler, in the upcoming months.

Next, we discuss the validation of and results from our initial simulation experiments.

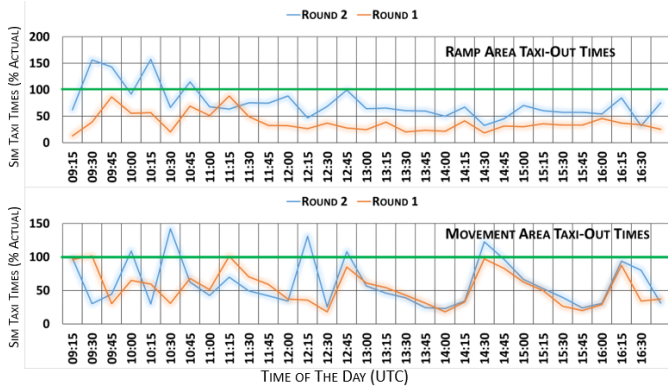
##### B. Simulation Validation

The first step we undertook after generating the simulation environment was to validate it against real-world historical operations data. For this purpose, we developed a set of streamlined tools and processes to compare key operational performance metrics from the simulation data with counterpart performance metrics from the same day of historical operations using end-to-end merged track data from NASA's Sherlock ATM data warehouse [9].

Our early simulation validation results showed that the SOSS simulation platform was significantly under-predicting the taxi out times, i.e., simulated taxi-out times were significantly smaller than the corresponding actual taxi out times. SOSS was also over-predicting taxi-in times by a big margin. There was also a significant mismatch between simulation and reality in terms of start times (i.e., the first point

of ASDE-X in actual surveillance data compared with simulated time of the flight reaching the same location), takeoff times, landing times and end times (i.e., the last point of ASDE-X in actual surveillance data compared with simulated time of the flight reaching the same location).

After the early validation step (which we call “Round 1” in this paper), we undertook a number of simulation enhancements in collaboration with NASA SOSS simulation experts [8]. In particular, we made the following enhancements to the SOSS simulation and our surface-airspace simulation platform: (1) we developed and applied a method for correcting erroneous Gate-out times in the historical data feeds that drove the SOSS simulations, (2) we verified and modified SOSS runway separation constraint settings as well as taxi speed modeling settings by comparing SOSS settings with counterpart metrics from actual operational data, (3) we activated more realistic pushback modeling features in SOSS, (4) we modified simulated arrival flight taxi paths to match more closely with real operations, (5) we developed an external taxi re-routing scheduler to simulate current day procedures for diverting taxiing arrival flights to hardstands while they are waiting for their gates to be vacated. After making these enhancements we re-performed the comparison of actual versus simulated performance metrics. We call this step the “Round 2” validation in this paper.



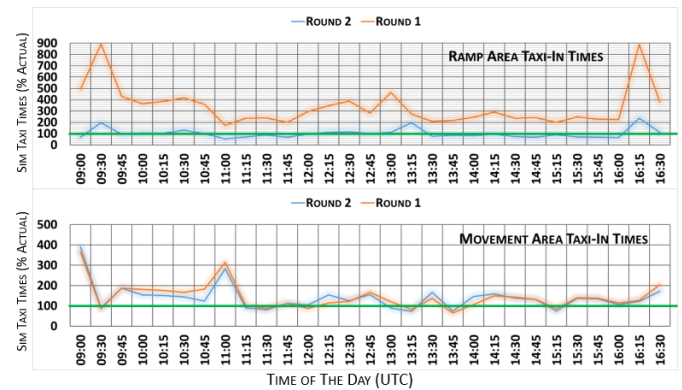
**Figure 4. Simulated Taxi-Out Times Shown as a Percentage of Actual Taxi-Out Times**

Our simulation modifications significantly improved the agreement between simulated and actual performance metrics. Figure 4 demonstrates the improvement in one key performance metric—taxi-out times. In this Figure, we plot the simulated taxi-out times per 15-minute time-bin (ramp times in the top-half, movement area times in the bottom-half). For each time-bin, simulated taxi times are shown as a percentage of the actual taxi-times (computed from historical ASDE-X data) for flights pushing back in the same time-bin. The ideal match with real operations data would be equivalent to the green line in both the plots (i.e., simulated times always equal to 100% of the actual times). We used an “area under the error curve” metric to obtain a good measure of how the actual and simulated taxi-out times matched with each other in Round 2 versus Round 1. To compute the “area under the error curve” metric we first calculate the area between the orange curve

(Round 1 simulated taxi-out times as a percentage of actual taxi-out times) and the green line (perfect match) in Figure 4. Then, we measure the area between the blue curve (Round 2 simulated taxi-out times as a percentage of actual taxi-out times) and the green curve. The difference between these two areas serves as a measure of improvement we obtained in Round 2 as compared to Round 1.

Round 2 simulations achieved significant improvement in matching the ramp area taxi-out times. In terms of the “area under the curve” metric, Round 2 provided a 70% improvement in the ramp area taxi times over Round 1. However, not much improvement was seen in movement area taxi-out times. For movement area taxi-out times the “area under the curve” metric was in fact 1% worse than Round 1, but we believe that difference is because of noise in the simulation data. Nevertheless, our simulation enhancements were not able to adequately address the discrepancy between simulated and actual movement area taxi-out times. In particular, a major source of discrepancy between simulation and reality was the following: we have observed in playback of real flight tracks that a lot of actual departures wait for a longer time than expected at the departure end of the runway before entering the runway or after entering the runway but before starting the departure roll. This wait may be because of the human delay in the local controllers issuing departure clearances. SOSS does not accurately simulate this additional wait time, which may be contributing to the movement area taxi-out time discrepancy. We plan to investigate this further to make additional improvements to the movement area taxi-out times.

Figure 5 demonstrates the taxi-in time improvement resulting from simulation modifications. As we did for the taxi-out times before, here we plot the simulated taxi-in times per 15-minute time-bin (ramp times in the top-half, movement area times in the bottom-half). Again, simulated taxi-in times are shown as a percentage of the actual taxi-in times for flights landing in the same time-bin. The ideal match would be equivalent to the green line in both the plots (i.e., simulated times are always equal to or 100% of the actual times).



**Figure 5. Simulated Taxi-In Times Shown as a Percentage of Actual Taxi-In Times**

We see from Figure 5 that Round 2 simulations achieved significant improvement in matching ramp area taxi-in times. There was very little improvement in movement area taxi-in times, but they were already very close to the actual movement area taxi-in times in Round 1. Further, computing the “area under the curve” metric, we see that Round 2 provided an 89% improvement in the ramp area taxi-in times and 6% improvement for movement area taxi-in times over Round 1.

### C. Simulation Results: Baseline vs. ATD-2 Comparison

As discussed above, the main purpose of the initial simulations presented here was to verify that the simulation environment is correctly modeling the impact of APREQ restrictions on the airport surface traffic and also the delay-alleviating impact of the ATD-2 scheduler. To this effect, in this section we present results comparing key performance metrics between baseline and ATD-2 simulations, for CLT departure flights that were scheduled to go through the constrained en route fixes (BEGVE and MALNR, as discussed above). We call these flights the “problem flights” from here onwards for the sake of brevity because these flights were impacted by the modeled APREQ constraints. We also discuss the comparative differences in the metrics between baseline and ATD-2 simulations computed over *all* the flights as a side-note, but the main focus is to verify the simulation of the impact of en route constraints on “problem flights”.

Figure 6 shows histograms of taxi-out times for the problem flights, for both current-day operations simulation (red bars) and ATD-2 operations simulation (green bars). As seen from the Figure, the taxi-out times are mostly similar across the two simulations with the ATD-2 simulations displaying slightly higher taxi-out times. The average taxi-out time for problem flights in the baseline simulation was 7.9 minutes as compared to 8.1 minutes in the ATD-2 simulation. Moreover, if averaged over all the flights in the simulation (a total of 266 departure flights), the taxi-out times for both baseline and ATD-2 simulations were both 8.9 minutes. This shows that the ATD-2 scheduler performed a good job of keeping the taxi-out times at a manageable level, while providing a big reduction in the airborne delays (as we will see later in this Section) by increasing adherence to APREQ windows.

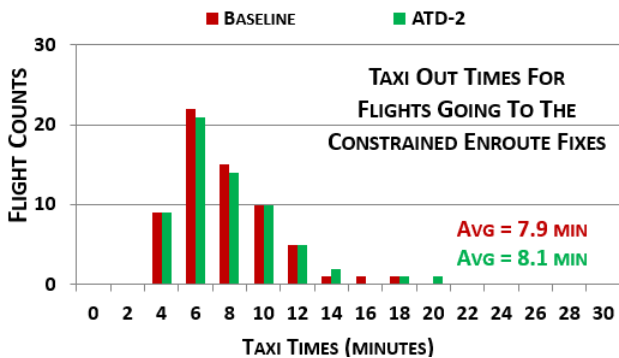


Figure 6. Simulated Taxi-Out Times for Problem Flights – Baseline (red) versus ATD-2 (green)

Figure 7 shows histograms of gate-holding delays (in minutes) for the problem flights, for both current-day operations simulation (red bars) and ATD-2 operations simulation (green bars). As seen from the Figure, the ATD-2 scheduler correctly holds more problem flights at their gates and releases them in time to make their APREQ takeoff time windows. Whereas, the baseline simulation only holds a small percentage of the problem flights at their gates resulting in lesser adherence to the APREQ window constraints and (as we will see later in this Section) higher en route airspace delays. The average gate-hold delay for problem flights in baseline simulation was 3.1 minutes as opposed to 4 minutes for the ATD-2 simulation. However, we also saw that the ATD-2 scheduler was delaying a lot of the non-problem flights at their gates as well with an average gate-hold delay of 3 minutes. This delay is introduced to balance the runway system demand to its capacity and may be counter-productive because it may starve the runway system at certain times. We are currently investigating applying a Target Departure Queue Length parameter to the scheduler which will get rid of unnecessary gate-hold delays that are starving the runway system. Moreover, the in-house version of ATD-2 scheduler performs runway/fix time-slot allocation as opposed to the NASA ATD-2 scheduler which applies pairwise separations between consecutive operations on the runway system. We also plan to fix this and other differences between the two schedulers in the near future.

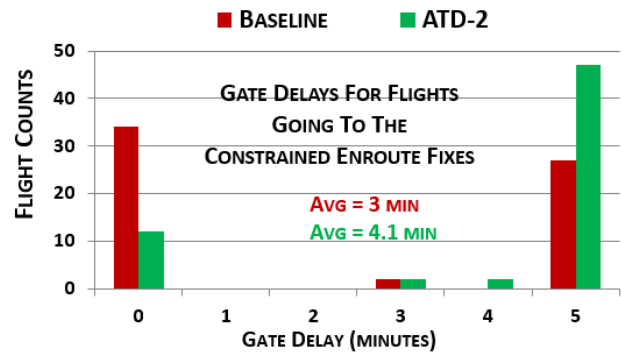


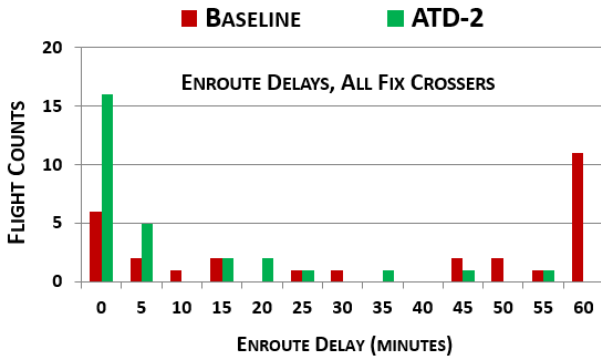
Figure 7. Gate-Hold Delays for Problem Flights

Finally, Figure 8 shows histograms of en route delays experienced by flights that crossed the constrained en route fixes (MALNR and BEGVE) before the end of the simulation. These are the delays experienced by CLT departures after crossing the departure fix but before merging into the en route traffic streams at the constrained en route fixes. As seen from the Figure, the ATD-2 scheduler was successful in saving a large amount of airborne delay by strategically holding constrained departure flights at their gates for the appropriate amount of time and releasing them just in time to make their APREQ window. As a result of larger adherence to APREQ windows in the ATD-2 simulation, most departure flights seamlessly merged into the overhead en route traffic streams with little or no airborne delay. The number of flights experiencing more than five minutes of airborne (en route) delay in the ATD-2 simulations was only 8 (out of a total of 29



en route fix crossers). Whereas, in the baseline simulation the number of flights with more than five minutes of airborne delay was 21 (out of the 29 fix crossers). This difference was because in the baseline simulation a large percentage of the flights were not able to meet their APREQ windows.

Note that the en route delays computed by our simulation are only an indicative metric. In the real world, as soon as the Center controllers or the System Command Center traffic managers start seeing en route airborne delays in the excess of 15-30 minutes, they will typically impose stricter APREQ or ground-hold restrictions on the departure airport. Our simulation does not model such second order effects and hence we see some excessive (> 30 minute) en route delays.



**Figure 8. Enroute Delays (in minutes) for APREQ Impacted Flights**

The total average delay (sum of gate, taxi and airborne delays) over all problem flights was equal to around 27 minutes in the baseline simulation, with a large portion (around 75%) of the total delay experienced in the air. Whereas, for the ATD-2 simulation the average total delay was around 18 minutes with approximately 50% of the total delay experienced on the airport surface (at the gate or in taxi).

This section discussed results from our initial simulation experiments. We plan to conduct more simulations using the chosen simulation days (see Table 1) at CLT airport as well as at the two other sites, which we have selected for detailed simulation-based assessments—EWR and DFW. Benefits assessments will be conducted using high-fidelity simulations of a small number of carefully chosen traffic scenarios at these three airports. Results from these simulations will be extrapolated to estimate ATD-2 benefits on a nation-wide and annualized scale. The next section describes our approach for this benefits extrapolation task.

## V. BENEFITS EXTRAPOLATION APPROACH

First, we describe our approach for nationalizing the ATD-2 benefits estimates

### A. Benefits Extrapolation to NAS-wide Airports

We propose to adopt a two-step approach to extrapolate the ATD-2 benefits to a NAS-wide benefit assessment. The first step will be a first-order estimate of implementing ATD-2 at different Core 30 airports, which will reflect the benefits of a

widespread deployment; the second step will consider the potential benefits that may be experienced NAS-wide due to the implementation of ATD-2 at a particular airport.

#### 1) Step 1: Extension to Core 30 Airports

This initial step will rely on the development of medium-fidelity queuing network models for the major airports. These models would focus on modeling the aggregate queuing behavior, at the runway thresholds, and if needed, the ramp areas. These models will consider major flows/configurations, and the number of departure runway servers needed. While such medium-fidelity modeling strategies will be recommended for the Core 30 airports, they will be developed and validated for a select subset of airports. The candidates for such validation are CLT, EWR, DFW, BOS, LGA, PHL [12-16]. In particular, the first three of these (CLT, EWR and DFW) will be conducted in tandem with the corresponding SOSS simulation models. We have developed queuing network models for both the north and south flows at CLT, and have commenced the model development for EWR. For some of the remaining airports, a simple single-queue model (representing the departure throughput) will be recommended. These models will help evaluate the first-order benefits of ATD-2 (especially as it relates to departure metering or improving departure throughput) to the performance of the airport at which the system is implemented.

#### 2) Step 2: Estimating NAS-wide Network Impacts

Implementation of ATD-2 at an airport is likely to yield benefits elsewhere in the system. From Step 1, we can see that by improving departure throughput and avoiding surface gridlock at an airport with ATD-2, the departure delays at that airport are likely to decrease. However, due to the interconnected nature of the system, this decrease in departure delays will imply less propagation of delays to other airports in the system, compared to a situation without ATD-2. Similarly, another ATD-2 benefit mechanism, improved merging and sequencing into the overhead stream, will also result in better on-time arrival performance at destination airports, thereby decreasing delay propagation. Our approach to estimating these network effects will leverage our recent work on modeling air traffic delay propagation [17-19]. These models reflect the effect of delays at one airport on future (i.e., over the next few hours) delays at other airports. We will employ these models to estimate the effect of decreasing the departure delays at one airport (e.g., CLT) through the implementation of ATD-2, on the delays at other Core 30 airports.

We propose to also compare this approach to an alternative approach that has been previously employed for the NAS-wide extrapolation of TFDM benefits [20]. This alternative methodology considered the share of TFDM benefits that were estimated at the subset of airports that are modeled in detail (e.g., CLT, EWR, and DFW), compared to the total benefits at the first-tier airports. This was estimated to be 14.1% [20]. Leveraging this finding, we can estimate the NAS-wide benefit of ATD-2 by multiplying the total benefits

estimated at these three airports from the SOSS/medium-fidelity simulations by an appropriately computed scaling factor.

### B. Benefits Extrapolation to Annualized Benefits

In order to extrapolate benefits to a full year, some metrics are required across the year. While the FAA Systems Operations Services' Office of Performance Analysis provides 20 "canonical" dates that are intended to be suitable for scaling to a full year with the simple ratio of 365/20, the high-fidelity surface-airspace simulation environment is not planned to be used to evaluate all 20, or at all major (Core 30) airports, thus additional information will be required. We apply the medium-fidelity queuing network models described in Section V.A to the analysis of ATD-2 benefits at a wider set of airports and dates. The identified limited set of target dates, provided in Table 1, will be run in the high-fidelity as well as medium-fidelity models, which allows for validation and calibration of the medium-fidelity models. After the medium-fidelity models are validated, all the 20 canonical dates will be simulated using these models and the NAS network level extrapolation applied which will potentially use a combination of other metrics (e.g., WITI, delay (taxi-out, departure, arrival)). At the annual level the benefits for the 20 days will be increased by a factor of 365/20 to create the baseline years' benefits. This benefit estimation method is based on guidance from the FAA.

## VI. CONCLUSIONS

A simulation-based approach has been developed for estimating the benefits of implementing the ATD-2 system for selected airport sites. Current-day operational shortfalls and relevant ATD-2 benefit mechanisms have been identified. A high-fidelity surface-airspace simulation environment has been developed to model operations at the selected airport sites including modeling of all the identified shortfalls and benefit mechanisms. Initial simulations show that the simulation environment correctly models the impact of APREQ TMIs on the airport surface traffic. They also revealed that enhancements are needed in certain areas, e.g., improving the agreement between the in-house ATD-2 scheduler model and NASA's official ATD-2 scheduler. A benefits extrapolation approach based on medium-fidelity airport models has been developed to extrapolate benefits results from few high-fidelity simulations to a nationwide and annualized scale.

### ACKNOWLEDGMENT

The authors would like to thank Rich Coppenbarger, Yoon Jung, and Robert Windhorst from NASA Ames Research Center and Zhifan Zhu and Sergei Gridnev from Stinger Ghaffarian Technologies for providing insights to support this study. We would also like to thank Michelle Eshow from NASA Ames Research Center for providing us access to NASA Sherlock ATM data warehouse to support our data analysis work. This research has been conducted under a NASA sponsored NRA grant.

## REFERENCES

- [1] NASA, "A Concept for Integrated Arrival/Departure/Surface (IADS) Traffic Management for the Metroplex," Airspace Technology Demonstration 2 (ATD-2) ConOps Synopsis (DRAFT), 2015.
- [2] Stroiney, S.R., and B. Levy, 2011. "Departure Queue Management Benefits Across Many Airports", 10th ICNS Conference, Herndon, VA.
- [3] Wood, Z., Kistler, M., Rathinam, S., and Jung, Y., "A Simulator for Modeling Aircraft Surface Operations at Airports," AIAA Modeling and Simulation Technologies Conference, AIAA 2009-5912, Chicago IL, 2009.
- [4] Windhorst, R., "Towards a Fast-time Simulation Analysis of Benefits of the Spot and Runway Departure Advisor," AIAA Guidance, Navigation, and Control Conference, AIAA 2012-4975, Minneapolis, MN, 2012.
- [5] Griffin, K., Saraf, A., Yu, P., Stroiney, S., Levy, B., Solveling, G., Clarke, J., and Windhorst, R., "Benefits Assessment of a Surface Traffic Management Concept at a Capacity-Constrained Airport," AIAA Aviation Technology, Integration, and Operations Conference, AIAA 2012-5533, Indianapolis, IN, 2012.
- [6] FAA, "FY2015 Sample Day Selection Results," Slides presented to the FAA ANG-B7 Systems Analysis & Modeling Division, March 29, 2016.
- [7] Klein, A., Kavoussi, S., Hickman, D., Simenauer, D., Phaneuf, M., & MacPhail, T. (2007, April). Using a Convective Weather Forecast Product to Predict Weather Impact on Air Traffic: Methodology and Comparison with Actual Data. In *Integrated Communications, Navigati* Klein, A., Kavoussi, S., Hickman, D., Simenauer, D., Phaneuf, M., & MacPhail, T. (2007, April). Using a Convective Weather Forecast Product to Predict Weather Impact on Air Traffic: Methodology and Comparison with Actual Data. In *Integrated Communications, Navigation and Surveillance Conference, 2007. ICNS'07* (pp. 1-10). IEEE.
- [8] ATAC Team for NASA NRA titled Benefit and Cost Assessment of Integrating Arrival, Departure, and Surface Operations with ATD-2, "Report Describing the Models, Simulation Environment, and Assumptions for Benefits Assessment at Select Sites," NASA NRA Project Report, July 2017.
- [9] Eshow M., Lui M., and Ranjan S., "Architecture and Capabilities of a Data Warehouse for ATM Research" Digital Avionics Systems Conference (DASC), 2014 IEEE/AIAA 33<sup>rd</sup>, Hilton Head, SC, 2014.
- [10] Windhorst, R.D., Montoya, J.V., Zhu, Z., Gridnev, S., Griffin, K.J., Saraf, A., and Stroiney, S., "Validation of Simulations of Airport Surface Traffic with the Surface Operations Simulator and Scheduler," AIAA Aviation Technology, Integration, and Operations Conference, Los Angeles, CA, August 2013.
- [11] Bagasol, L., "ATD-2 Tactical Scheduler," NASA ATD-2 Tactical Scheduler description slides, June 2017.
- [12] S. Badrinath and H. Balakrishnan. "Control of a non-stationary tandem queue model of the airport surface," American Control Conference, May 2017.
- [13] I. Simaiakis and H. Balakrishnan. "A Queuing Model of the Airport Departure Process," *Transportation Science*, Vol. 50, No. 1, pp. 94-109, February 2016.
- [14] I. Simaiakis, M. Sandberg and H. Balakrishnan. "Dynamic Control of Airport Departures: Algorithm Development and Field Evaluation," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 1, pp. 285-295, February 2014.
- [15] P. McFarlane and H. Balakrishnan. "Optimal Control of Airport Pushbacks in the Presence of Uncertainties," American Control Conference, July 2016.
- [16] I. Simaiakis and H. Balakrishnan. "Design and Simulation of Airport Congestion Control Algorithms," American Control Conference, June 2014.
- [17] K. Gopalakrishnan, H. Balakrishnan and R. Jordan. "Deconstructing Delay Dynamics: An Air Traffic Delay Example," International Conference on Research in Air Transportation (ICRAT), June 2016.

- [18] K. Gopalakrishnan, H. Balakrishnan and R. Jordan. "Stability of Networked Systems with Switching Topologies," IEEE Conference on Decision and Control, December 2016.
- [19] K. Gopalakrishnan and H. Balakrishnan. "A Comparative Analysis of Models for Predicting Delays in Air Traffic Networks," USA/Europe Air Traffic Management Seminar, June 2017.
- [20] FAA, "Final Business Case for Terminal Flight Data Manager," FAA document, June 2016. FAA, "Final Business Case for Terminal Flight Data Manager," FAA document, June 2016.