Metrics to Characterize Airport Operational Performance Using Surface Surveillance Data

Harshad Khadilkar

PhD Candidate

Department of Aeronautics and Astronautics

Massachusetts Institute of Technology, Cambridge, MA 02139

harshadk@mit.edu

Hamsa Balakrishnan

Associate Professor

Department of Aeronautics and Astronautics

Massachusetts Institute of Technology, Cambridge, MA 02139

hamsa@mit.edu

Harshad Khadilkar, Hamsa Balakrishnan

Abstract

Detailed surface surveillance datasets from sources such as the Airport Surface Detection Equipment, Model-X (ASDE-X) have the potential to be used for the analysis of airport operations, in addition to their primary purpose of enhancing safety. This paper describes how surface surveillance data can be used to measure airport performance characteristics in three different ways: (1) Characterization of surface flows, including identification of congestion hotspots, queue dynamics and departure throughput; (2) Development of metrics to evaluate the daily operational performance; and (3) Development of metrics to gauge long-term performance across different runway configurations and operating conditions. The proposed metrics have been developed with active feedback from operations personnel at Boston Logan International Airport, and are therefore evaluated and discussed using this airport as an example. These metrics can provide useful feedback on operational performance to airport operators, and therefore have the potential to improve the efficiency of surface operations at airports.

Harshad Khadilkar, Hamsa Balakrishnan

Metrics to Characterize Airport Operational Performance Using Surface Surveillance Data

Harshad Khadilkar, Hamsa Balakrishnan

NOMENCLATURE

ADS-B	Automatic Dependent Surveillance - Broadcast
ARTCC	Air Route Traffic Control Center
ASDE-X	Airport Surface Detection Equipment, Model-X
ASPM	Aviation System Performance Metrics database
ASQP	Airline Service Quality Performance database
ATCT	Airport Traffic Control Tower
BOS	Boston General Edward Lawrence Logan International Airport
EDCT	Expected Departure Clearance Time
FAA	Federal Aviation Administration
GPS	Global Positioning System
η	Departure Spacing Efficiency
NAS	National Airspace System
TFM	Traffic Flow Management
ZBW	Boston ARTCC

I. INTRODUCTION

I.A Motivation

Airports form the critical nodes of the air transportation network, and their performance is a key driver of the capacity of the system as a whole [1], [2], [3]. With several major airports operating close to their capacity for large parts of the day [4], the smooth and efficient operation of airports has become essential for the efficient functioning of the air transportation system. Operational efficiency of airports is influenced to a great extent by the characteristics of surface movement of aircraft taxiing out for departure, or taxiing to their gates after landing.

Largely driven by the availability of data, studies of airport operations have traditionally focused on airline operations and on the aggregate estimation of airport capacity envelopes [5] [6] [7] [8]. Most of this research is based on data from a combination of the Aviation System Performance Metrics (ASPM) [9] and the Airline Service Quality Performance (ASQP) databases. These databases provide the times at which flights push-back from their gates, their takeoff and landing times, and the gate-in times, as reported by the airlines. ASPM also provides airport-level aggregate data, including records of runway configuration use and counts of the

Harshad Khadilkar, Hamsa Balakrishnan

total number of arrivals and departures in 15 min increments. Such data can be used to develop queuing models of airport operations [10] [11], [12] or empirically estimate airport capacity envelopes [7], [8]. However, the level of detail in these datasets has been insufficient to investigate other factors that affect surface operations, such as interactions between taxiing aircraft, runway occupancy times, queue sizes on the airport surface, etc. The recent deployment of Airport Surface Detection Equipment, Model-X (ASDE-X) at major airports has made it possible to continuously track each aircraft on the surface, and has thereby enabled the analysis of surface operations in greater detail than ever before.

This paper proposes, for the first time, techniques by which ASDE-X data can be leveraged to characterize airport surface operations, and to develop metrics for formally measuring day-to-day as well as long-term airport operational performance. These metrics have been developed with the active support of the air traffic controllers and the operations manager at Boston Logan International Airport, and therefore address those areas of airport performance that are most relevant to operational personnel.

The purpose of this work is twofold. Firstly, it emphasizes the hitherto unexplored value of using surface surveillance data for post-hoc performance analysis, in addition to its primary intent as a real-time safety tool. If the vast repository of ASDE-X data is fully exploited, important insights can be obtained regarding the characteristics of an airport.

Some indications of these characteristics already exist in the form of anecdotal rules of thumb. It is now possible to corroborate or refute these ideas based on empirical evidence.

The second purpose of this work is to propose a set of standardized and generic performance measures for airports. Since ASDE-X systems are now installed at several major airports in the United States, these metrics can be used to make comparisons across airports. The algorithms described in this paper are fully automated: once the raw data is obtained, a shell script executes all tasks including data extraction, parsing, analysis, and result storage, without supervision. The proposed analysis methods and operational metrics are illustrated for the example of Boston Logan International Airport (BOS), based on data from August 2010 to May 2012. Daily performance plots of the type presented in this paper have been compiled by the authors, and shared with the Operations Manager at the Boston Airport Traffic Control Tower (ATCT) on a regular basis since November 2010.

I.B Overview of ASDE-X data

ASDE-X is primarily a safety tool designed to improve situational awareness and to mitigate the risk of runway collisions [13]. It incorporates real-time tracking of aircraft on the surface to detect potential conflicts. There is potential, however, to use the data generated by it for the analysis of surface operations and airport performance, and for the prediction of quantities such as taxi times. ASDE-X data is generated by sensor fusion from three primary sources: (i) Surface movement radar, (ii) Multilateration using onboard transponders, and (iii) GPS-enabled Automatic Dependent Surveillance Broadcast (ADS-B). Reported parameters include each aircraft's position, velocity, altitude and heading. The update rate is about 1 Hz for each individual track. This study uses ASDE-X data from Boston Logan International Airport (BOS), as collected by the MIT Lincoln Laboratory's Runway Status Lights (RWSL) system. This system was first installed at Boston in late 2009, and a regular data feed has been available to MIT starting from mid-2010.

Departing flights are tracked starting from the time when the onboard transponder is turned on: while the timing of this event varies from airport to airport, analysis shows that at BOS, transponder capture typically occurs while the flight is still in the ramp area, on average about 5 min after the aircraft is cleared for push-back [14]. The push-back clearance at Boston Logan is given by the air traffic controller for all aircraft apart from those in Terminal A. Aircraft start their engines after push-back is completed, and typically start their transponders when calling for taxi clearance. At Terminal A, the ramp towers controls the aircraft up to the spot, which is when control is passed to the ATC Tower. However, transponders are still started as soon as the aircraft starts to taxi within the ramp. The ASDE-X coverage extends approximately 18 nmi out from the airport, which is when the flight track for departures ends. All the results presented in this work are based on ASDE-X data, both because of its superior detail and also to ensure a valid comparison between the different performance metrics.

I.C Data analysis methods

In order to parse ASDE-X data into measurable quantities, several analysis algorithms were developed. The high level of detail in these data results in large file sizes, with each day's data from Boston Logan occupying approximately 2 gigabytes of space. This requires the development of highly efficient methods for analysis. Moreover, raw ASDE-X tracks contain a substantial amount of noise as well as exogenous hits such as ground vehicles and inactive aircraft being towed on the surface. Additional complications are added by the existence of multiple flights with the same call-sign and aircraft tail number during the duration of a day's operations. Therefore, the data needs to be preprocessed before it can be used for analysis.

Mitigation of the effects of noise is done by a multimodal unscented Kalman filter that produces smoother estimates of aircraft position, velocity and heading [15]. The filtering algorithm uses models of aircraft dynamics in order to correct for intermittent errors in the flight tracks. Such errors can leak into the raw data from the fusion algorithm which handles three separate data sources, or from radar detection issues with aircraft at very low speeds. Additional error handling is carried out by a second algorithm, which includes (1) separation of continuing flights with the same call-sign, (2) detection and tagging of off-nominal operations such as cancelled takeoffs, go-arounds and aircraft absorbing Traffic Flow Management (TFM) related delay in active movement areas, and (3) removal of irrelevant tracks such as ground vehicles and helicopters using dedicated and separate routing. The combined filtering technique significantly improves the reliability of the detected flight tracks, which are then transferred to various analysis algorithms. Each flight track is tagged with a departure/arrival time and runway. By tracking each aircraft from push-back to wheels-off, various airport states such as the active runway configuration, location and size of the departure queue, and departure/arrival counts are measured. In addition, airport-level performance metrics such as average taxi-out times, runway usage and departure spacing statistics are also tracked. The definitions of these metrics and the algorithms proposed for measuring them are described in subsequent sections of this paper.

II. CHARACTERIZATION OF SURFACE OPERATIONS

II.A Departure queue characteristics

Visualization of the filtered ASDE-X data yields insights into the dynamics of airport operations. It can help identify locations on the surface where aircraft typically queue up for departure, for different runway configurations and operational procedures. For example, Figure 1

shows the layout of BOS with the different runways. Figure 2 shows the departure queue (aircraft icons) formed on September 09, 2010 in the 22L, 27 | 22R configuration, i.e., when Runways 22L and 27 (on the east side) were being used for arrivals and 22R (on the west side) was being used for departures.

Figure 1: Layout of Boston Logan International Airport.

Figure 2: Visualization of queuing behavior at BOS. Aircraft icons represent departing aircraft queuing for departure from Runway 22R. The current queue size is 12 aircraft, with 7 aircraft on the east side and 5 aircraft on the west side of the runway.

In Figure 2, the departure queue can be seen forming at the threshold to runway 22R. In this figure, some aircraft are seen in a separate queue on the west side of the runway, waiting for crossing clearance. This is due to construction on the taxiway to the west of the runway during 2010. Such subtleties could be observed by using the ASDE-X data to generate animations in Google Earth®. Observation of operations over several days of data was used to identify typical queue formation areas for each configuration at Boston. These areas were then designated in the analysis codes for automatic tracking of departure queues and calculation of statistics such as time spent by individual aircraft in the queue and the variation of queue length over the course of each day. Note that each departing aircraft was tagged as being in queue if it was within the designated box for its departure runway, and below a certain threshold velocity. Some runways required the definition of more than one queuing area.

Harshad Khadilkar, Hamsa Balakrishnan

It was found that this separation of multiple queues was instructive. For example, aircraft that are absorbing delay on the surface due to TFM initiatives usually occupy the same queuing area, while flights without delay occupy another. Therefore, high occupancy of the former queue is typically indicative of bad weather in the NAS. Figure 3 shows the mean time spent in the departure queue as a function of the queue length, based on data from all runway configurations. The queue length is defined as the number of aircraft in the departure queue as seen by a new aircraft just joining it. It can be seen that on average, an additional aircraft in queue entails a penalty of 83 s for each of the aircraft behind it. This value seems reasonable when one considers standard departure separations, as discussed in Section III. The standard deviation of the time in queue increases from approximately 20 sec for the lower queue sizes, to approximately 60 sec for the larger queue sizes. It is likely that the value is somewhat enhanced by the possibility of swaps within the queue, which could disproportionately increase or decrease the time spent in queue for some of the aircraft.
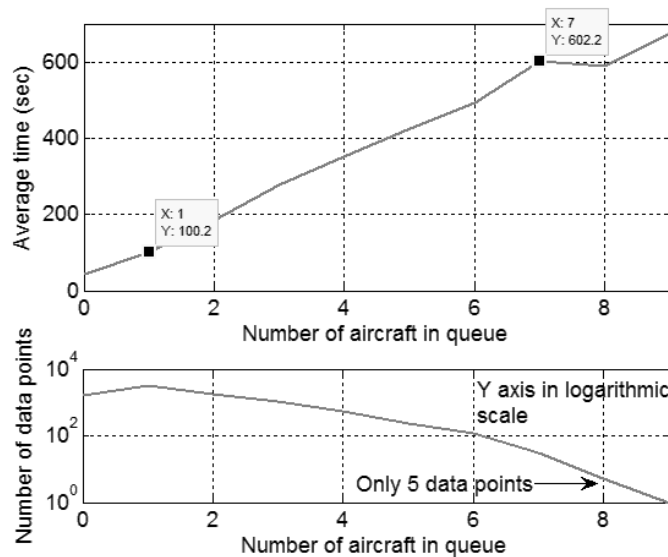

Figure 3: Time spent in departure queue as a function of queue length when the aircraft enters it.

II.B Departure throughput characteristics

Departure queue characterization offers an insight into surface operations from an aircraft's perspective. In order to analyze airport-level operational performance, the variation of departure throughput (defined as the number of takeoffs from the airport in a 15 min interval) with the number of active departing aircraft on the surface is considered. An aircraft is defined to be active from the time of first transponder capture (first detection with ASDE-X) until its wheels-off time. Previous studies have shown using ASPM data that the departure throughput increases with the addition of aircraft to the surface, until the maximum sustainable throughput is reached, and the airport surface saturates [11], [4]. This observation is further corroborated by the results produced using ASDE-X data. Figure 4 shows the departure throughput curve for a specific configuration at BOS. Only aircraft with jet engines are counted in this analysis, because propeller-driven aircraft are fanned out via separate departure fixes at BOS, and do not affect the jet departure process [14]. The throughput curves for other configurations are similar in nature to

Harshad Khadilkar, Hamsa Balakrishnan

Figure 4, differing only in the point of saturation and corresponding maximum throughput. These differences can be attributed to configuration-specific procedures, such as closely spaced departures on intersecting runways.
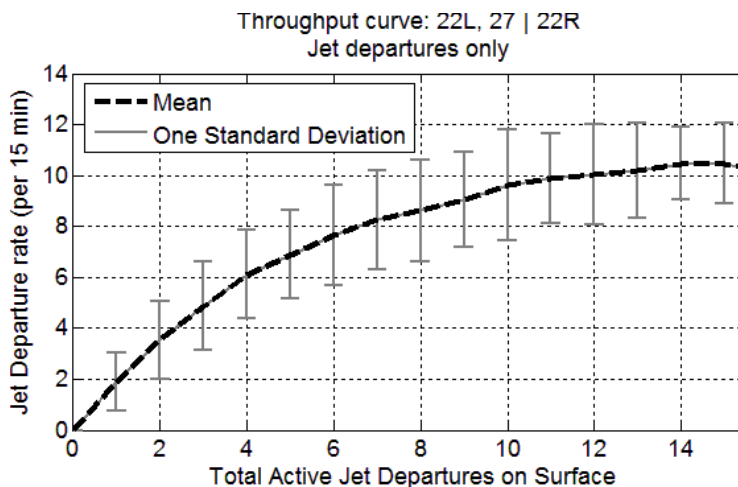

Figure 4: Departure throughput as a function of the number of active departing aircraft on the surface.

III. METRICS TO CHARACTERIZE DAILY OPERATIONAL PERFORMANCE

Building upon the results described in the previous section, metrics to measure the daily operational performance of an airport are presented here. The objective is not to evaluate individual air traffic controller performance, but to look for systemic inefficiencies and to identify opportunities for improvement. While each airport has particular rules, regulations and procedures that must be considered in order to be consistent across different configurations and time periods, the basic concepts used to define these metrics can be generalized to any airport. For example, there are subtle variations in the strategies used by different airports to accommodate both arrivals and departures on the same runway or on intersecting runways. At BOS, arrivals have to cross the departure runway in the 22L, 27 | 22R configuration, while in the 27 | 33L configuration, it is the departures that have to cross the arrival runway. Arrivals are controlled by the Boston TRACON or the ZBW Air Route Traffic Control Center, and not the Air Traffic Control Tower (ATCT). Inter-arrival separations and arrival sequences are not under the control of the tower, and are not indicative of ATCT performance.

This paper proposes three metrics of day-to-day operational performance that account for practical complexities while keeping the computational effort at a reasonable level. These metrics are important for identifying the effects of off-nominal operations, which can be lost by looking at only aggregate-level data.

III.A Average taxi-out times

The most natural performance measure from the point of view of passengers is the average taxi-out time for departures. The taxi-out time is defined in this paper to be the difference between the time of first transponder capture and the wheels-off time. At Boston Logan, aircraft transponders are usually turned on just before taxi clearance is given by ATC. Therefore, while the transponder capture time might not correspond exactly to the push-back time, it is a good measure of the time taken for actual taxi.

Harshad Khadilkar, Hamsa Balakrishnan

The taxi-out time is an important quantity that affects not only flight delays but also taxi-out fuel consumption [16]. In general, taxi-out times are highest during the peak congestion periods. At BOS, these are the morning departure push between 0600 h and 0800 h local and the evening push between 1900 h and 2000 h local time. Figure 5 shows the variation of average taxi-out times on a sample day. The averages are calculated over 15 min intervals for the entire day. Each bar in the upper plot represents the average taxi-out time experienced by the aircraft pushing back in that 15 min interval. The number of pushbacks in the corresponding interval is shown in the lower plot. The peaks in both pushbacks and taxi-out times around 0700 h and 1900 h can be seen clearly.

Note that in the calculation of this metric, long delays absorbed by flights on the surface have been removed. These aircraft usually have specified departure times known as Expected Departure Clearance Times (EDCTs), decided by constraints elsewhere in the National Airspace System (NAS). While it is desirable to have aircraft absorb these delays at the gate, it is not always possible because of conflicts with arrivals that are scheduled to park at the same gate. In such cases, aircraft at Boston Logan absorb the delay in a separate designated area (called the Juliet pad), or in a separate runway queue as explained in Section II A. Figure 6 shows a plot similar to the one in Figure 5, but with flights with EDCTs included in the data. A comparison of the two plots illustrates the effect of excluding these flights. Section IV.B will discuss how the average taxi-out time metric can be extended to assess the long-term performance of an airport.
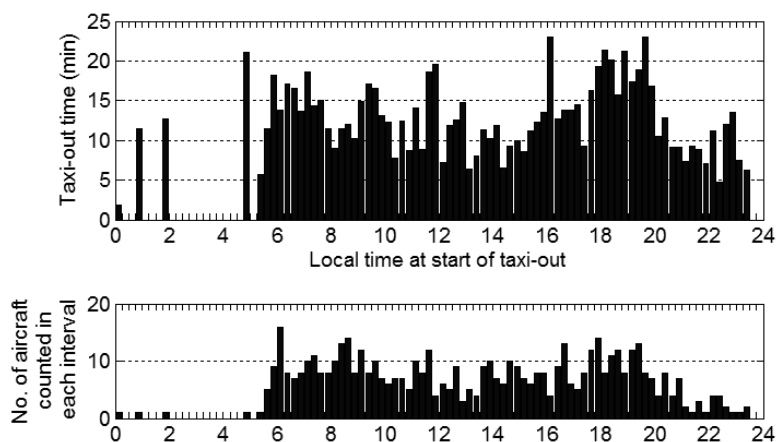

Figure 5: Average taxi-out times by time-of-day at BOS on Dec. 09 2010, with excessive hold times removed.
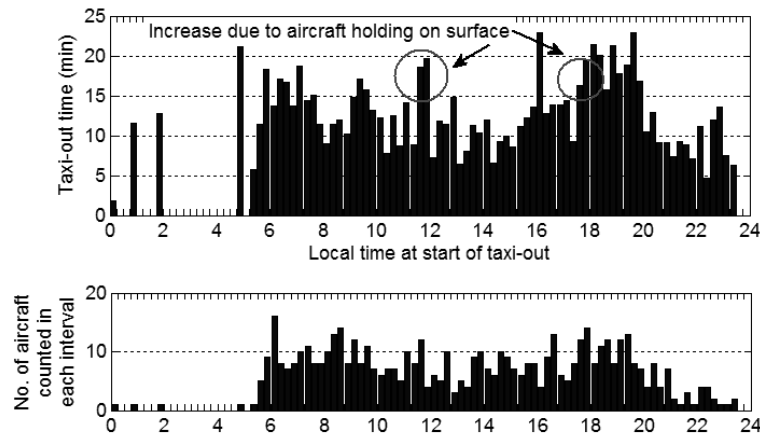
Harshad Khadilkar, Hamsa Balakrishnan

Figure 6: Average taxi-out times by time-of-day at BOS on Dec. 09 2010, with all hold times included.

III.B Runway utilization

The most capacity constrained element in departure operations is the runway, as shown in [2]. Therefore, it is important to ensure that an airport's runway system is used as efficiently as possible. To look at the current usage characteristics of runways at BOS, a metric called Runway Utilization was defined. The utilization is expressed as a percentage, calculated for every 15 min interval. It is given by the fraction of time in the 15 min interval for which a particular runway is being used for active operations. The types of active operations that are accounted for in calculating the runway utilization are as follows:

1. *Departure:* An aircraft on the runway, between the start of its takeoff roll and wheels-off.
2. *Hold:* A departing aircraft holding stationary on the runway, waiting for takeoff clearance.
3. *Approach:* Counted from the time an aircraft is on short final (within 2.5 nmi of runway threshold) to the time of touchdown.
4. *Arrival:* Counted from the moment of touchdown to the time when the aircraft leaves the runway.
5. *Crossings/Taxi:* Counted when aircraft are either crossing an active runway, or taxiing on an inactive runway.

Harshad Khadilkar, Hamsa Balakrishnan

Analysis algorithms detected each of these operating modes by using the filtered states from ASDE-X tracks for each aircraft. The 'approach' phase was included in the runway utilization because no other operations can be carried out on a runway when an aircraft is on short final. Even though, technically, the aircraft is not on the runway, ignoring this operational constraint would give an erroneously low utilization figure for an arrival runway. The approach areas corresponding to runways 15R/33L and 9/27 at Boston are shown in Figure 7. Since the ASDE-X system is tuned for surface surveillance, altitude information far away from the airport is highly unreliable. Therefore, the approach areas shown in the figure are two-dimensional and do not include an altitude restriction. This is not a severe limitation, because high-altitude overflights are not captured by ASDE-X. False alarms (aircraft that are within the approach areas but are not landing at the airport) in the designated areas are thus largely non-existent.

Figure 7: Approach areas for two runways at BOS. Aircraft within the white cones are assumed to be utilizing the runway.

Figure 8 and Figure 9 show the utilization plots for a sample good-weather day, for three different runways. The topmost plot in each figure shows the breakup of utilization for the runway over the course of the day. The plot in the middle shows operational counts in each 15 minute interval, for both ends of the runway. Finally, the lowermost plot shows the variation of queue length over the day.

It should be noted that the queue length is calculated for every second, while the top two plots are aggregate figures over the 15 min interval. Details such as configuration changes can be seen immediately from the utilization plots. For example, it can be seen from Figure 9 that departure operations shifted from runway 22R to runway 33L at 1000 h. The departure runway utilization (Figure 9 [Top]) was nearly 100% from 1800 h to 2000 h (apart from two notable instances which are discussed below), the period of peak evening demand. The split in utilization between the approach mode and arrival mode in Figure 8 shows that aircraft spend a roughly equal amount of time on final approach, as compared to the actual touchdown and roll-out. The formation of a runway crossing queue can also be seen during the peak evening period. This is a common feature at Boston Logan when a combination of runways 33L and 27 are in use.

The value of acknowledging the coupling between operation counts, utilization and queue dynamics can be seen in Figure 9 [Top]. The presence of two heavy arrivals on runway 33L can be noted in the evening period between 1800 and 2000 h local. This is a common request at

Harshad Khadilkar, Hamsa Balakrishnan

Boston Logan because of the longer length available on runway 33L. The disruption caused by these events is clearly seen in the utilization plots, where the figure drops to approximately 60% in both cases. The arrival at 1830 h also causes a large drop in the number of departures in that time interval. It also has an effect on the average inter-departure spacing achieved, as discussed in the next section.
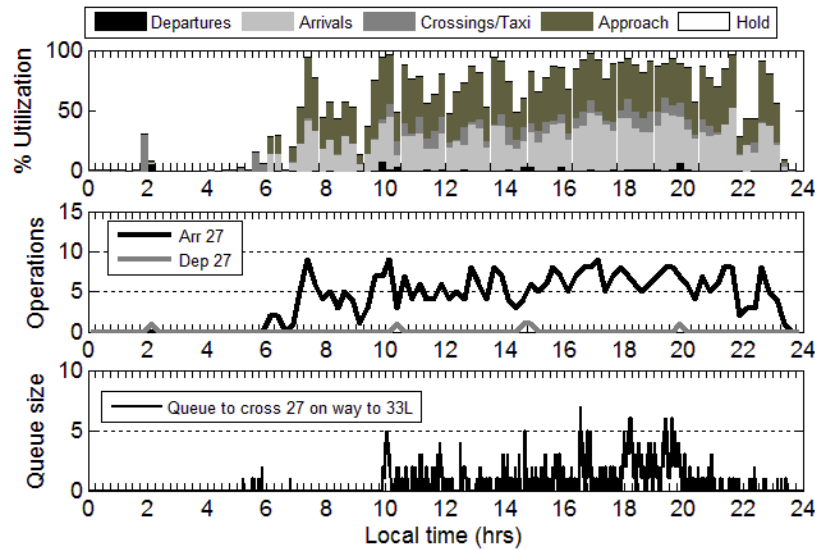

Figure 8: Utilization of Runway 9/27 on December 09, 2010. Runway 27 was used for arrivals for the entire day. An occasional departure can be seen in the second plot, accompanied by a dip in the number of arrivals and the utilization for the time period. In the bottom plot, queue formation can be seen, composed of aircraft waiting to cross the runway for departure on 33L.

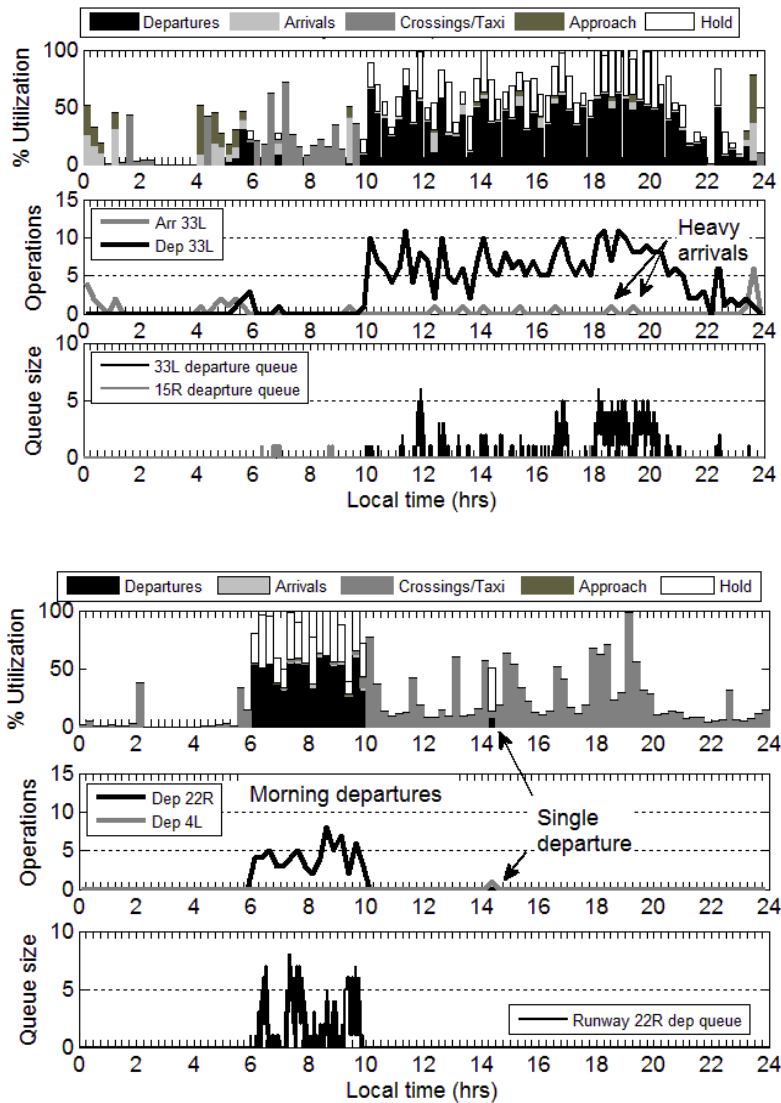Harshad Khadilkar, Hamsa Balakrishnan

Figure 9: Utilization of [Top] Runway 15R/33L and [Bottom] Runway 4L/22R on December 09, 2010. Runway 33L was inactive until 1000 h, after which it was used for departures. Occasional arrivals seen in the top figure are Heavy aircraft that request this runway due to its greater length. Runway 22R (bottom figure) was used for departures during the peak morning period. Longer departure queue lengths are seen here than for Runway 33L, due to two closely spaced departure banks at 0600 h and 0745 h.

By contrast, Figure 10 and Figure 11 show the runway utilization on October 27, 2012, when the 4L, 4R | 9, 4R configuration was active. This was one day prior to the arrival of Hurricane Sandy at Boston, and the evening was affected by bad weather. It can be seen that while operational counts and runway utilization appears normal in the morning (total departure rate across runways 9 and 4R of 10 per 15 minutes), both values are significantly lower in the evening.
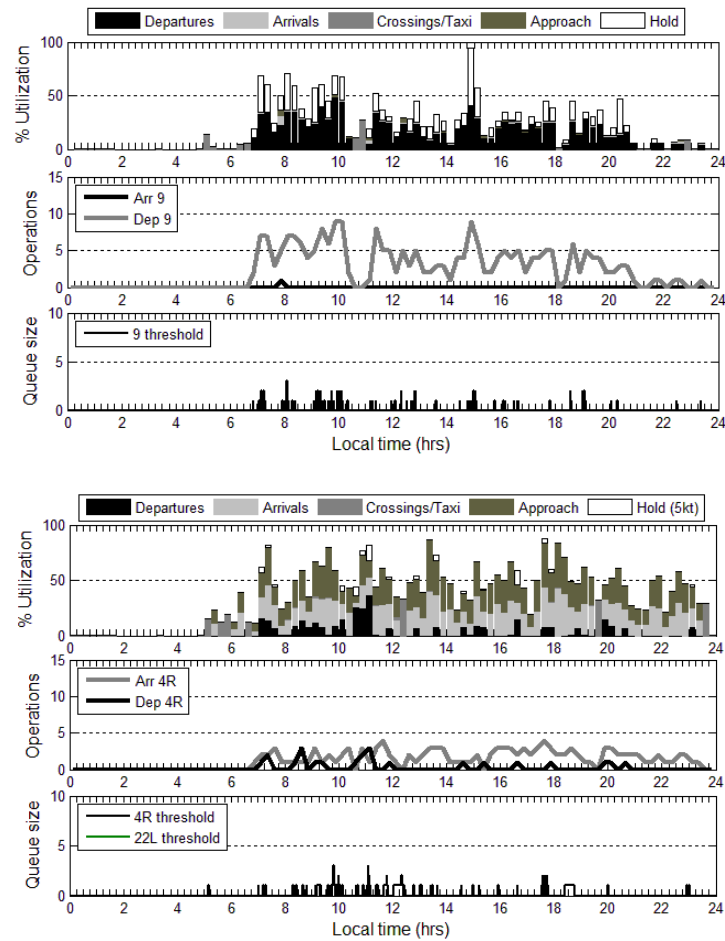
Harshad Khadilkar, Hamsa Balakrishnan

Figure 10: Utilization of [Top] Runway 9/27 and [Bottom] 4R/22L on October 27, 2012. On this day, the runway configuration is in the opposite direction to that of December 09, 2010. Runway 9 was used for departures throughout the day. Mixed departure and arrival operations are seen on Runway 4R, which is requested by Heavy aircraft because of its greater length.
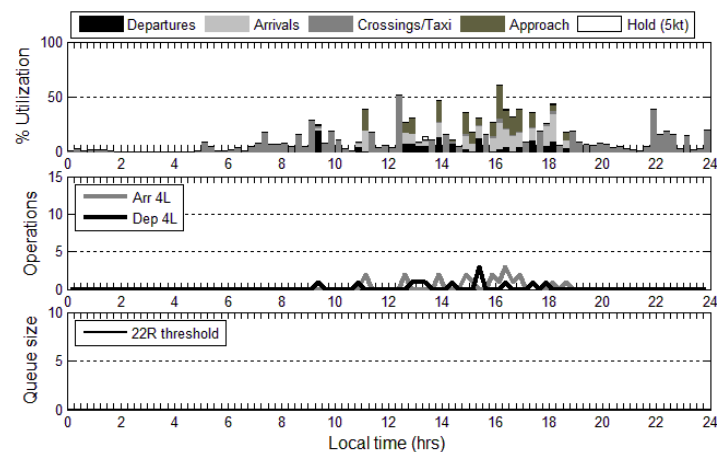

Figure 11: Utilization of Runway 4L/22R on October 27, 2012. Mixed departure and arrival operations are seen through the day, but only props are allowed to use Runway 4L for departures.

Ideally, it is desirable for the utilization value to be 100% for all active runways in times of peak demand. The sample figures show that while this value is achieved for much of the peak period, it is difficult to sustain. Disruptions may be caused by off-nominal events such as runway closures due to foreign objects, arrivals requesting a departure runway for landing, or gaps in the arrival sequence. It should be noted that the utilization for a departure runway is always higher than that for an arrival runway. This is because the departures can be packed close together, with the next aircraft in queue holding on the runway while the previous aircraft starts its climb-out. On the other hand, tightly packed arrivals would increase the risk of frequent go-arounds caused by aircraft not being clear of the runway quickly enough to allow the next arrival to land. Therefore, the arrival stream has a buffer in addition to the minimum spacing that is dictated by FAA regulations.

III.C Departure spacing efficiency

As noted earlier, departures can be spaced with a smaller safety buffer as compared to arrivals. However, the target departure spacing is still governed by a set of standards, customized to each airport depending on the runway and airspace layout. It is generally recommended to maintain a minimum spacing of 120 s for a departure following a heavy aircraft [17]. At BOS, the target separations based on a combination of regulatory requirements such as these, rules of thumb followed by the controllers, and average performance as measured using ASDE-X data, are as shown in Table 1.

Table 1: Target departure separations. The columns correspond to the weight class of the trailing aircraft, while the rows correspond to the weight class of the leading aircraft (P = Props, S = Small jets, L = Large jets, 757 = Boeing 757, H = Heavy jets). All figures are in seconds.

(s)	P	S	L	757	H
P	40	40	40	40	40
S	40	65	65	65	65
L	40	65	65	65	65
757	110	110	110	110	110
H	120	120	120	120	120

To compare the actual inter-departure separation with these target values, a metric called the Departure Spacing Efficiency was defined. As with runway utilization, this metric is calculated for each 15 min interval. However, the Departure Spacing Efficiency is not runway-specific, but addresses departure operations at the airport as a whole. To calculate it, the difference between the wheels-up times of each pair of consecutive departures is compared to the target level of separation for that pair, based on the aircraft classes of the leading and trailing aircraft. Each additional second more than the target level is counted as a second *lost*. Time is counted as 'lost' only if there are other aircraft in queue, waiting for departure. This ensures that the efficiency figure does not fall simply because of low demand. Note that Miles-in-Trail (MIT) restrictions may also result in additional inter-departure separations, which will be captured by this metric. Inopportune separations between arriving aircraft on a crossing runway can also cause a drop in departure efficiency. On the other hand, controllers can also sometimes manage to depart aircraft with a separation less than the target level, depending on factors such as the availability of multiple runways for departure. In this case, each second less than the target

Harshad Khadilkar, Hamsa Balakrishnan

separation level is counted as a second *gained*. Then, the Departure Spacing Efficiency, denoted η, in each 15 min interval is given by

$$\eta = 1.0 + \frac{Total\ seconds\ gained\ in\ 15\ mins - Total\ seconds\ lost\ in\ 15\ mins}{15 \times 60\ seconds}$$

It should be noted that the use of multiple runways does not always allow departures to take place with less than the target level of spacing. For example, at BOS, departure operations on runways 22R and 22L have to take place as on a single runway, because both sets of departures have to use the same departure fixes. However, when runways 4R and 9 are used for departures, aircraft can be spaced more closely, thus boosting the airport's efficiency [14]. Figure 12 demonstrates the calculation procedure for counting the time lost in a 15 min interval. The local time is shown on the x-axis. Each spike denotes the wheels-off time for a departure, with the height of the spike corresponding to the weight class of the aircraft. The spike then tapers off, reaching the 'clear to release' line when the target separation interval elapses. The gap from this point to the next departure spike counts towards the total number of seconds lost in the current 15 min interval.
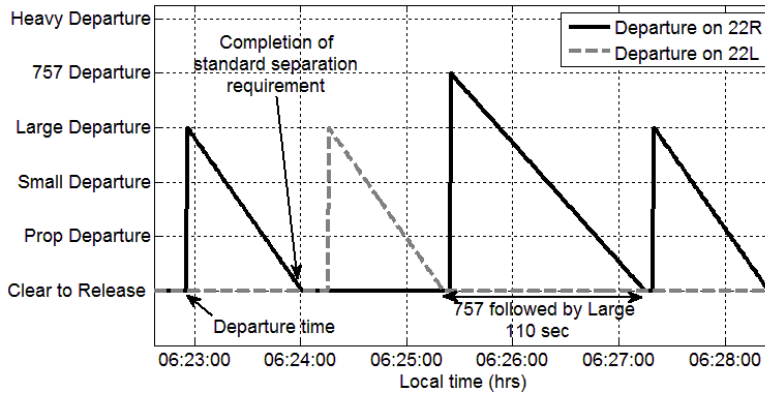

Figure 12: Visualization of consecutive departures on runways 22R and 22L.

Accounting for the effect of arrivals

There is, however, a caveat associated with calculation of the total time lost. As described previously, arrival spacing is not under the discretion of Boston Tower. In configurations where arrivals take place on a runway that is the same as or that intersects the departure runway, this can cause a dip in the efficiency. This effect is accounted for by discounting the idle time of a departure runway when an arrival is on short final (within 2.5 nmi of threshold) to an intersecting runway or to the same runway. Figure 13 and Figure 14 show the variation of Departure Spacing Efficiency with the time of day for December 09, 2010, with and without accounting for the arrival effect, respectively.

In each figure, the plot on the top shows the efficiency in each 15 min interval, while the plot on the bottom shows the departure count in the corresponding intervals. The colored bars in the middle indicate the departure congestion level at the airport, calculated using a combination of the departure counts and queue lengths. Note that there are several time periods (for example, between 1630 and 2015 hours), when not including the impact of arrivals would lead to the erroneous conclusion that the efficiency was lower than it actually was. A comparison of Figure

Harshad Khadilkar, Hamsa Balakrishnan

13 and Figure 14 suggests that the efficiency during this time was above 85% when accounting for arrivals, whereas it was as low as 75% when their effect was ignored.

Note the large dip in efficiency just prior to the configuration change at 1000 hours. On the other hand, a few intervals with net efficiency more than 1.0 can also be seen. These intervals correspond to spikes in the departure count, since consistent separation values less than the target level result in a large number of departures. The most notable high-efficiency interval is the one from 1945 to 2000 h, which is in the middle of a period with high demand. The bottom plot shows that the controllers managed to serve ten departures in this interval (nine on Runway 33L and one on Runway 27), while a comparison with Figure 8 shows that seven arrivals were also achieved. In this way, a combination of different performance metrics offers insights into the intricacies of surface operations that result in the net operational counts, which are the traditional measure of airport performance.
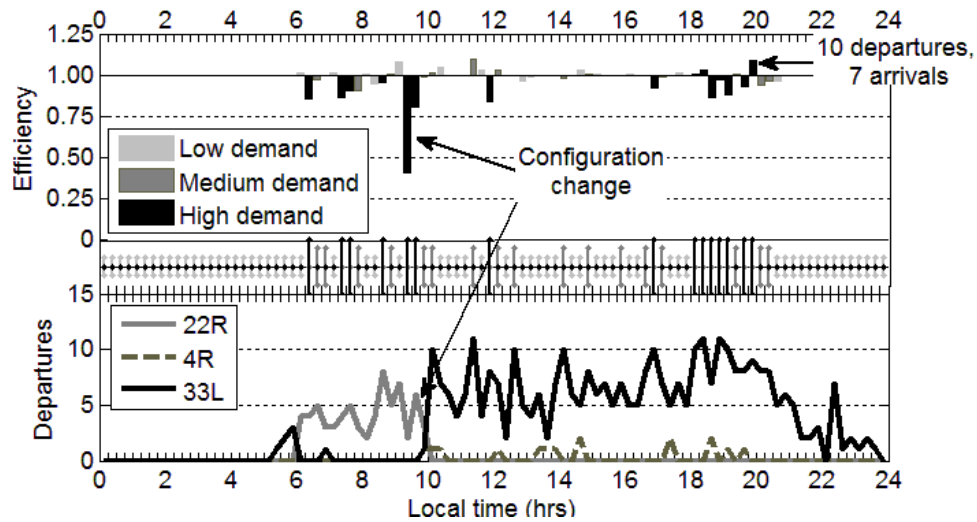

Figure 13: Departure spacing efficiency on December 09, 2010, accounting for the effect of arrivals on the same/crossing runway.
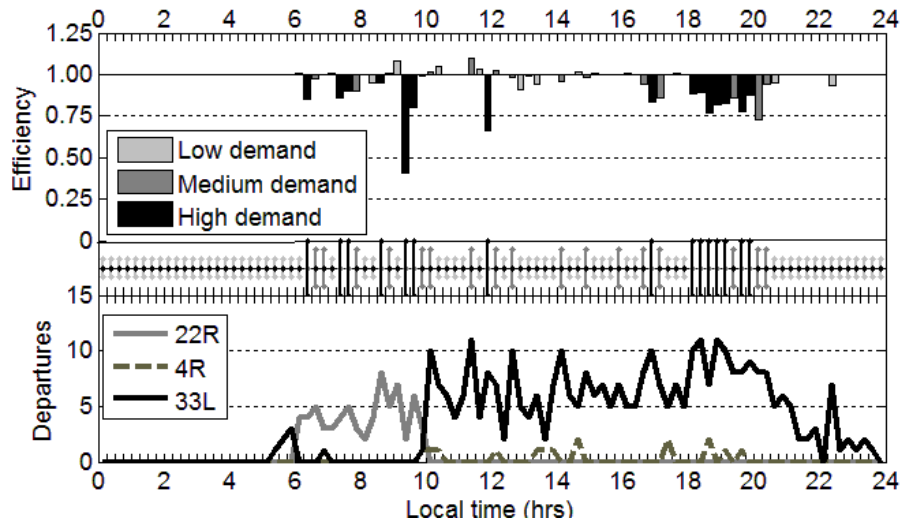
Harshad Khadilkar, Hamsa Balakrishnan

Figure 14: Departure spacing efficiency on December 09, 2010, not accounting for the effect of arrivals on the same/crossing runway.

IV. METRICS TO CHARACTERIZE LONG-TERM PERFORMANCE

IV.A Long-term average departure spacing efficiency

The metrics defined above, when consistently tracked over several months, can be used to measure the average operational performance of the airport. For example, Figure 15 shows the average departure spacing efficiency at BOS, sorted by configuration and demand levels. It is evident from the figure that for most configurations, the efficiency drops as demand increases. This conclusion is intuitive, since high demand usually means more operational complexity, more runway crossings, etc.

It is also noted that the airport is most efficient when departures are taking place from Runways 4R and 9. As mentioned before, this configuration allows closely-spaced departures on the two crossing runways, which enhances the efficiency. An interesting observation is that the 33L | 27 configuration, with departures on Runway 27, is more efficient than the 27 | 33L configuration, even though the two are operationally very similar. One possible explanation for the increased departure spacing seen on Runway 33L is that it is longer than Runway 27. As a result, a departure from 33L takes longer to be completely clear of the runway, which is typically the cue for the air traffic controller to release the next aircraft. Either of these configurations can be used when winds are from the northwest. The operational implication of this result is that it is more desirable to use the 33L | 27 configuration when the departure demand is high, particularly in the mornings when the arrival demand is relatively low.
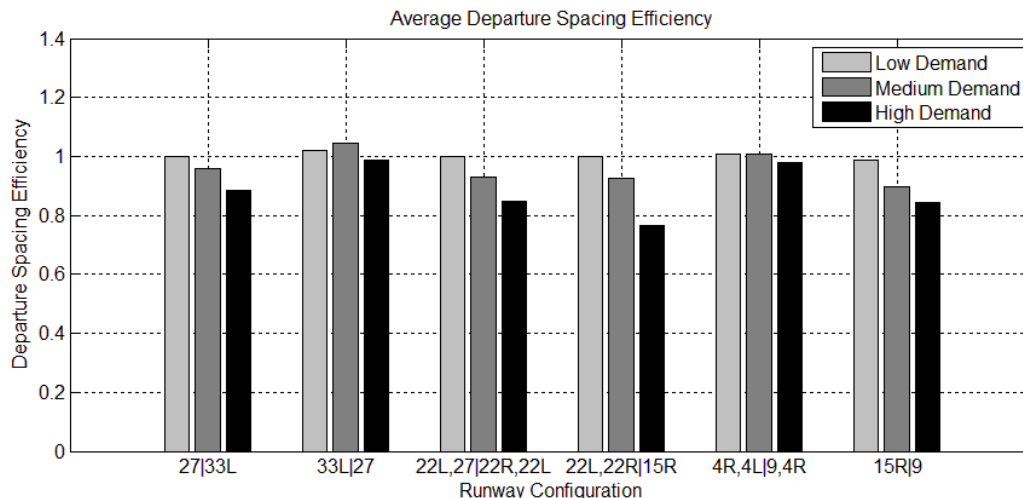
Harshad Khadilkar, Hamsa Balakrishnan

Figure 15: Average departure spacing efficiency for frequently used configurations at BOS.

IV.B Taxi-out time comparisons with historical data

Beyond providing feedback on the daily variation of 15 min averaged taxi-out times, it is also possible to compare these times with the historical average. The historical average taxi-out time over the previous three months of operations is parameterized by the current configuration of the airport, and the 'congestion level' as described in Section III.C. Figure 16 shows a sample plot from December 09, 2010, the same day as in Figure 2. It is seen that average taxi-out times followed the historical variation through most of the day, but that there were some intervals of high taxi-out times during the evening peak period. The 27 | 33L configuration that was active at this time (Figure 8) is one of the lower efficiency configurations, as was seen in Figure 15.

In addition, as seen in Figure 13, the spike in taxi-out times in the interval from 1600-1615 h corresponds to a 'moderate traffic' interval in the middle of generally low traffic intervals. No historical value is shown for intervals in which there were too few operations to define a configuration, or in which the airport was operating in a non-standard configuration.
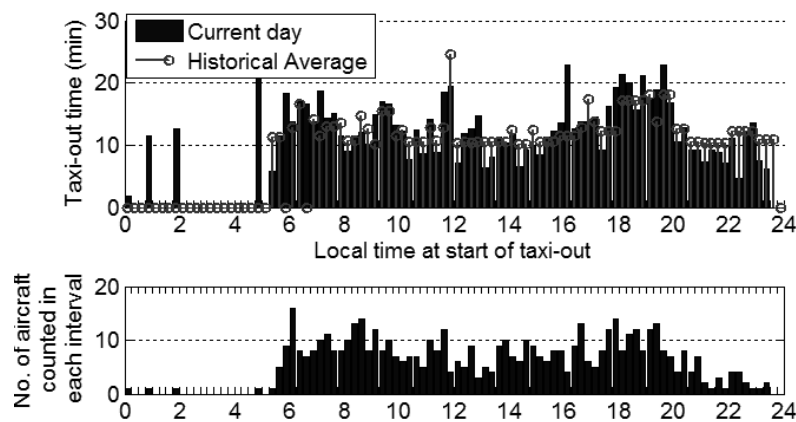

Figure 16: Variation of average taxi-out times by time-of-day on December 09, 2010.

Harshad Khadilkar, Hamsa Balakrishnan

V.   CONCLUSIONS

This paper presented several novel ways in which surface surveillance data can be used for the detailed analysis of airport surface operations. In addition to the qualitative assessment of surface operations through visualization, the paper showed how to directly estimate quantities such as departure queue statistics and airport departure throughput characteristics. It also proposed three metrics to quantify day-to-day airport performance, namely, average taxi-out times, runway utilization and departure spacing efficiency.

Metrics of long-term airport performance based on taxi-out times and departure spacing efficiency were also presented. The proposed metrics were discussed in detail for the case of Boston Logan International Airport, and it was shown that they could be used to gain insights into the performance of the airport. These insights can help to identify opportunities for the improvement of operational efficiency. For example, the results presented in Section IV.A showed that one of two symmetric configurations resulted in higher inter-departure spacing efficiency than the other. This fact could be leveraged to utilize the more efficient configuration when the choice was available.

Other examples of salient feedback include an analysis of the different methods used by controllers for active runway crossing. If done correctly, arriving aircraft can cross the departure runway while the next departing aircraft is lining up for takeoff, significantly improving runway utilization figures. Different methods are effective to different degrees depending on the airport and runway configuration, and only a detailed analysis of empirical data can reveal the extent of their success.

The assessment techniques proposed in this paper can be easily extended to the analysis of other airports, as has been demonstrated for the case of New York's LaGuardia (LGA) and Philadelphia (PHL) airports in recent work [18]. In the future, monitoring algorithms can be developed to automatically flag off-nominal events in real-time, and display notifications to air traffic controllers. This research effort would need to include human factors studies as well as data mining algorithms for event identification. The fuel burn and emissions impact of any congestion control algorithms implemented at the airport can also be usefully analysed using the methods proposed in this paper, and this aspect is addressed in [14], [16].

ACKNOWLEDGMENTS

REFERENCES

[1]   H. R. Idris, B. Delcaire, I. Anagnostakis, W. D. Hall, N. Pujet, E. Feron, R. J. Hansman, J.-P. Clarke and A. Odoni, "Identification of flow constraint and control points in departure operations at airport systems," in *AIAA Guidance, Navigation and Control Conference*,

Harshad Khadilkar, Hamsa Balakrishnan

1998.

[2]   R. de Neufville and A. Odoni, Airport Systems: Planning, Design, and Management, McGraw-Hill, 2003.

[3]   N. Pyrgiotis, K. Malone and A. Odoni, "Modelling delay propagation within an airport network," *Transportation Research Part C,* pp. 60-75, 2013.

[4]   I. Simaiakis and H. Balakrishnan, "Impact of Congestion on Taxi Times, Fuel Burn and Emissions at Major Airports," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2184, 2010.

[5]   E. Pels, P. Nijkamp and P. Rietveld, "Inefficiencies and scale economies of European airport operations," *Transportation Research Part E,* vol. 39, no. 5, pp. 341-361, 2003.

[6]   D. Gillen and A. Lall, "Developing measures of airport productivity and performance: an application of data envelopment analysis," *Transportation Research Part E,* vol. 33, no. 4, pp. 261-273, 1997.

[7]   E. Gilbo, "Airport capacity: representation, estimation, optimization," *IEEE Transactions on Control Systems Technology,* vol. 1, no. 3, pp. 144-154, 1993.

[8]   V. Ramanujam and H. Balakrishnan, "Estimation of Arrival-Departure Capacity Tradeoffs in Multi-Airport Systems," in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009.

[9]   Federal Aviation Organization, "Aviation System Performance Metrics Database," [Online]. Available: http://aspm.faa.gov/main/aspm.asp.

[10] N. Pujet, B. Delcaire and E. Feron, "Input-output modeling and control of the departure process of congested airports," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Portland, OR, 1999.

[11] I. Simaiakis and H. Balakrishnan, "Queuing Models of Airport Departure Processes for Emissions Reduction," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2009.

[12] H. Idris, J. Clarke, R. Bhuva and L. Kang, "Queueing model for taxi-out time estimation," *Air Traffic Control Quarterly,* vol. 10, pp. 1-22, 2002.

Harshad Khadilkar, Hamsa Balakrishnan

[13] Sensis Corporation, "ASDE-X Brochure - ASDE-X 10/05.qxd," East Syracuse, NY, 2008.

[14] I. Simaiakis, H. Khadilkar, H. Balakrishnan, T. G. Reynolds, R. J. Hansman, B. Reilly and S. Urlass, "Demonstration of Reduced Airport Congestion Through Pushback Rate Control," in *Ninth USA/Europe Air Traffic Management Research and Development Seminar*, 2011.

[15] H. Khadilkar and H. Balakrishnan, "A multi-modal Unscented Kalman Filter for inference of aircraft position and taxi mode from surface surveillance data," in *11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, 2011.

[16] H. Khadilkar and H. Balakrishnan, "Estimation of aircraft taxi fuel consumption using flight data recorder archives," *Transportation Research Part D,* vol. 17, pp. 532-537, 2012.

[17] Federal Aviation Administration, "Pilot and Air Traffic Controller Guide to Wake Turbulence".

[18] M. Sandberg, "Applications of ASDE-X Data to the Analysis of Airport Surface Operations," Masters Thesis, Massachusetts Institute of Technology, May 2012.