DEPARTURE TAXI TIME PREDICTIONS USING ASDE-X SURVEILLANCE DATA

Jeffrey Legge, Benjamin Levy
Sensis Corporation

Keywords: ASDE-X, ATL, Departure, Taxi Time, Prediction

Abstract

Accurate prediction of departure taxi times will help airlines to proactively manage push-times, to optimally assign scarce ramp resources, and to propagate delay information to destination airports in a more timely fashion. Air Traffic Control (ATC) will benefit via improved demand forecasts for the terminal area and en-route air sectors. An ancillary benefit to such predictions is the ability to discern factors contributing to longer taxi times.

To facilitate accurate predictions, we will analyze the utility of the Airport Surface Detection Equipment, Model X (ASDE-X) surface surveillance data. Months of archived data support both historical analysis (i.e., under similar conditions, what happened in the past?) and a more dynamic real-time surface analysis (e.g., aircraft positions, queues, and runway utilization).

We will present results based on data from March 2008 at Hartsfield – Jackson Atlanta International Airport (ATL). However, all analyses could be equally applied to any ASDE-X equipped airport.

1 Introduction

A primary responsibility of airport surface management is to orchestrate the ongoing procession of taxiing aircraft between gates and runways. To make this procession as efficient as possible, an accurate estimate of taxi time is essential.

For departures, an accurate estimate of taxi time may be used to identify the optimal pushback time with respect to minimizing taxi-time. Given the present fuel crisis facing the airline industry, finding a method to shave minutes off taxi times will save airlines millions of dollars per year. [1] Looking to the future, value will be achieved via a reduction in the emission of hydrocarbons, carbon monoxide, and nitrous oxide. [2]

An accurate estimate of taxi time will also result in more accurate Estimated Time of Departure (ETD) reports. These reports may be utilized by ATC for load balancing both on the airport surface and in en-route air sectors. A more accurate ETD will also mitigate the largest source of error in Estimated Time of Arrival (ETA) reports. [3, 4]

In the past, many analyses predicting taxi times have relied on a combination of static data determined far before the event (e.g., scheduled out time) and a tabulation of events that have already occurred (e.g., actual out and off times for other aircraft). In particular, these analyses have not had access to a real-time surveillance source (e.g., ASDE-X).

The ASDE-X surveillance system provides 1 Hz surveillance for all surface targets with approximately 8m xy-resolution. This level of granularity allows for dynamic analyses of surface queues and runway utilizations, creating potential for more accurate ETDs.

Our research to-date has been dominated by the mundane but necessary tasks of parsing, sanitizing, and integrating disparate data sources. In this paper, we briefly summarize the data sources and data routines used. Subsequently, we present results for a baseline ETD algorithm (no knowledge), a present-day ETD algorithm (no surface surveillance knowledge), and then examine how much additional value real-time surface surveillance could provide.
2 Data Sources

Several data sources were correlated together both to support the analysis and, when possible, to provide independent validation.

2.1 Airport Map

The map data used for this research consists of three layers:

- An AutoCAD DXF drawing provides a subdivision of the surface into regions. These are used exclusively as a background image to help orient the data.
- Runway regions are derived from the Digital Aeronautical Flight Information File (DAFIF). These are used to identify operations from the surveillance data.
- Ramp areas are hand drawn to capture the origination ramp area for each departure. A frequency distribution of the initial surveillance report for all departures is used to determine the general ramp areas.

Figure 1 illustrates the map data at ATL, including the initial surveillance report overlay.

2.2 ASDE-X Surveillance

The ASDE-X surveillance data provides position updates grouped into tracks. These tracks are post-processed to produce operations (i.e., arrivals and departures). For this research, we consider the month of March 2008. During this month, our algorithms detected 41,546 arrivals and 41,599 departures. The imbalance of +53 departures is not unreasonable and indicates that our arrival and departure detection algorithms are roughly equal in accuracy.

Figures 2 and 3 illustrate the surface utilization for arrivals and departures, respectively. These figures give a feel for the surveillance produced by the ASDE-X system.

The shading in Figures 2 and 3 is calculated as follows:

\[ f_{ij} = \log(cellCount_{ij} + 1) \]  

(1)

---

\(^1\) The ASDE-X system at ATL was not designed to provide surveillance in the ramp (non-movement) areas. Although surveillance is still available, it is of a lower quality than that in the movement area. Thus, it is not presently possible to consistently and accurately determine the precise gate from which each aircraft departs.
In Equation 1, a logarithm is used to modify the reported frequency counts \( c_{ij} \) for each 10x10 meter cell. Without this scaling step, the output is limited to ramp areas and runway queues (areas where the most time is spent). In Equation 2, a value \( c_{ij} \) between zero and one is produced which is then linearly mapped to a grayscale.

The cell shading figures also serve as a basis for evaluating the frequency with which aircraft utilize specific surface routes. This can be achieved by conditioning the results based on the origination ramp area and/or the destination runway.

The ASDE-X track data also consists of (when available) Mode-S ID and flight ID (callsign).

2.3 FAA Aircraft Registry

The FAA aircraft registry, available as a free download [5], is used to derive aircraft tail numbers from the Mode-S IDs reported by ASDE-X (for Mode-S equipped aircraft). The aircraft registry may also be used to determine aircraft type for the purposes of fuel burn, emissions analyses, or wake-vortex spacing. [1]

2.4 ASPM

Aviation System Performance Metrics (ASPM), available for download to registered users via the FAA [6], provides operational counts, airport configuration, and airport weather. The highest resolution for this data is 15 minute blocks.

To validate the ASDE-X data, we binned the ASDE-X operations into 15 minute blocks and compared with the “Efficiency Computation” numbers from ASPM. These numbers are compiled based on flights that actually flew. Figure 4 illustrates the difference between the ASDE-X and ASPM counts in March 2008.

![ASDE-X vs. ASPM Operation Counts](image)

Fig. 4. ASDE-X vs. ASPM Operation Counts

The roughly symmetric distribution with a mean near zero is expected – the ASDE-X and ASPM data show comparable arrivals and departures. A closer analysis indicates that the ASPM sees more operations than ASDE-X.

### Table 1. ASDE-X vs. ASPM Operation Counts

<table>
<thead>
<tr>
<th></th>
<th>ASDE-X</th>
<th>ASPM</th>
<th>PCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZ</td>
<td>41599</td>
<td>42276</td>
<td>97.4</td>
</tr>
<tr>
<td>DZ</td>
<td>41546</td>
<td>42713</td>
<td>98.3</td>
</tr>
</tbody>
</table>

This data indicates that our present operation detection algorithms are not perfect – on average our operation detection algorithms undercount 25.5 arrivals and 37.2 departures per day, as compared with the ASPM data. However, this data also indicates that the ASPM data may not be perfect – it is unlikely that 437 fewer aircraft were at the airport at midnight 2008-04-01 compared with midnight 2008-03-01.

2.5 BTS

The Bureau of Transportation Statistics (BTS) freely provides “On-Time Performance” data. [7] This data contains the Out-time, Off-time, On-time, and In-time (OOOI) for aircraft flown by US certified air carriers that account for at least one percent of domestic scheduled passenger revenues.

We correlate this data with the ASDE-X data by comparing tail numbers and off-times. As noted earlier, the ASDE-X coverage is not as good in the ramp area. However, by associating
the ASDE-X and BTS data, we obtain a more reliable estimate of the out-time. It is noted that there is no standard for reporting OOOI times; so, it is likely that there are variations by carrier as well as human errors.

For the month of March 2008, BTS reports 35,387 departures, or approximately 85.2% of those reported by ASDE-X. This is reasonable considering that BTS does not report all operations. Of these 35,387 departures, we are able to match 33,501, or approximately 94.7%.

2.6 Other

Additional data sources parsed but not used in this paper include Aviation Routine Weather Reports (METARs) and Aircraft Situation Display to Industry (ASDI) data.

METARs provide more comprehensive weather information than that provided by ASPM. This data is generally updated once per hour unless unusual conditions necessitate otherwise. ASDI data provides flight plan information including first fix. In our later results, we assume that first fix information combined with the current airport configuration can be used to derive the specific runway a departure will use.

3 Results

Our results are split into three sections. First, we demonstrate the complexity of the problem by considering the simple prediction of the mean taxi time for all data. Then, we demonstrate the gains achieved by conditioning on qualitative attributes such as airline, meteorological condition, origination ramp area, and destination runway. Finally, we consider how much extra value might be provided by a real-time surveillance system such as ASDE-X.

3.1 Baseline ETD

Figure 5 illustrates the distribution of taxi-out times seen at ATL during the month of March 2008.

![Figure 5. Taxi-Out Time Distribution](image)

This data is for the 33,501 operations matched between ASDE-X and BTS. Each data point is calculated as ASDE-X off-time minus the associated BTS out-time.

The data in Figure 5 has a sample mean of 19.9 minutes and a sample standard deviation of 9.6 minutes. The minimum value is 1.9 minutes, and the maximum value is 180.4 minutes. The mode is 14 minutes.

The 99th percentile of the data is 54.4 minutes. The 336 data points greater than this value are considered outliers and are discarded from further analysis. In the future, such data points will need to be reliably identified based on available real-time information.

Excluding outliers, the remaining 33,165 data points in Figure 5 have a sample mean of 19.3 minutes and a sample standard deviation of 7.3 minutes.

By predicting a taxi-out time of 19.3 minutes for all aircraft, we can calculate the error as predicted taxi-out time minus actual taxi-out time. This distribution looks identical to Figure 5 except that it is right-shifted by the mean (19.3 minutes). We use the sample standard deviation (still 7.3 minutes) as a metric for how good our prediction is.

3.2 Conditioned ETD

A slightly more sophisticated approach is to categorize departures by commonly available parameters (e.g., origination ramp area) and to make a taxi-out time prediction on a per-category basis. This approach will be most
useful when there is a parameter (or set of parameters) that explains a significant portion of the variance present in the taxi-out time distribution. To reduce the impact of statistical artifacts, categories with membership under 1000 (approximately three percent of the data) are coalesced into a generic “Other” category.

In Tables 2-5, we present results for the parameters: origination ramp area, destination runway, airline, and meteorological condition. In the case of ramp areas, the labels correspond to the map in Figure 1 as follows:

- The northern row of six adjacent areas is labeled 1-6 from left to right.
- The southern row of six adjacent areas is labeled 7-12 from left to right.
- The remaining areas are labeled 13-19 counterclockwise from the upper left.

Table 2. Taxi-Out Time by Ramp Area

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5757</td>
<td>18.5</td>
</tr>
<tr>
<td>10</td>
<td>4427</td>
<td>19.6</td>
</tr>
<tr>
<td>3</td>
<td>4167</td>
<td>18.0</td>
</tr>
<tr>
<td>9</td>
<td>3795</td>
<td>19.0</td>
</tr>
<tr>
<td>8</td>
<td>3042</td>
<td>20.5</td>
</tr>
<tr>
<td>2</td>
<td>2666</td>
<td>20.3</td>
</tr>
<tr>
<td>5</td>
<td>2259</td>
<td>18.4</td>
</tr>
<tr>
<td>11</td>
<td>1998</td>
<td>20.6</td>
</tr>
<tr>
<td>1</td>
<td>1861</td>
<td>19.6</td>
</tr>
<tr>
<td>7</td>
<td>1797</td>
<td>19.9</td>
</tr>
<tr>
<td>Other</td>
<td>1396</td>
<td>20.9</td>
</tr>
<tr>
<td>Total</td>
<td>33165</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Taxi-Out Time by Runway

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>26L</td>
<td>11757</td>
<td>19.1</td>
</tr>
<tr>
<td>27R</td>
<td>8835</td>
<td>19.8</td>
</tr>
<tr>
<td>8R</td>
<td>6104</td>
<td>18.1</td>
</tr>
<tr>
<td>9L</td>
<td>5469</td>
<td>19.1</td>
</tr>
<tr>
<td>Other</td>
<td>1000</td>
<td>26.2</td>
</tr>
<tr>
<td>Total</td>
<td>33165</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Taxi-Out Time by Airline

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAL</td>
<td>11999</td>
<td>20.1</td>
</tr>
</tbody>
</table>

Table 5. Taxi-Out Time by Meteorological Condition

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMC</td>
<td>26366</td>
<td>19.0</td>
</tr>
<tr>
<td>IMC</td>
<td>6799</td>
<td>20.5</td>
</tr>
<tr>
<td>Total</td>
<td>33165</td>
<td></td>
</tr>
</tbody>
</table>

The pair-wise results (computed but not shown) were computed and similarly do not significantly reduce the sample standard deviation. In particular, knowledge of the origination gate area and the destination runway is not sufficient to eliminate the variability present in the data.

### 3.3 Real-time Surveillance Aided ETD

The inability of simple explanatory variables to account for much variability in the data suggests that the problem may require a more dynamic solution. We investigate the utility of the virtual departure queue, defined as follows: The virtual departure queue (VDQ) is defined for a specific aircraft A and a specific time T as the number of aircraft that depart from runway R between T and T' where R is the runway A eventually departs from at time T'.

Figure 6 shows the relationship between VDQ and taxi-out time. The solid red horizontal line is the cutoff for outliers – data points above the line were discarded in earlier analyses. The dotted red line models equality between VDQ and taxi-out time (departures at a rate of one per minute).
As expected, there is a clear linear relationship between VDQ and minimum taxi-out time. This relationship degrades for smaller values of VDQ. At this point, the minimum taxi-out time seems to be roughly constant and is likely dominated by the minimum theoretical taxi-out time (approximately 7-8 minutes).

To compare with the results from Section 3.2, we fit the data with the optimal linear solution:

\[
\text{taxiOutTime} = 0.92 \times \text{VDQ} + 9.8
\]  

(3)

This algorithm lowers the standard deviation of the predicted minus the actual taxi-out times from 7.3 minutes to 5.1 minutes – a substantial improvement compared to the conditioning in Section 3.2. However, unlike the simple algorithms presented in Section 3.2, this algorithm relies on data that is not available in real-time.

To make this algorithm real-time, we need a method to predict VDQ. A very simple algorithm to predict VDQ is to simply count the number of aircraft with the same destination runway that have pushed back before the aircraft of interest but which have not yet departed. This data is illustrated in Figure 7.

The optimal linear solution for the data in Figure 7 is:

\[
\text{taxiOutTime} = 0.65 \times pVDQ + 12.6
\]  

(4)

The standard deviation achieved via this algorithm is 6.5 minutes. As expected, this is not as strong as the relationship between taxi-out time and VDQ; however, it is our contention that the simple algorithm presented here for predicting the virtual queue may be improved by leveraging the situational knowledge provided by ASDE-X.

4 Summary

We developed a rich data set to investigate the taxi-out time for departures at ATL in the month of March 2008. Several methods for predicting taxi-out time were considered including conditioning of the data and measuring the virtual departure queue. Standard deviation of the prediction error was used as a metric to evaluate each method. The results from Sections 3.1, 3.2, and 3.3 are summarized in Table 6. The VDQ entry in Table 6 is starred because it is not derived based on real-time data.

<table>
<thead>
<tr>
<th>Sample Std. Dev.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditioned</td>
<td>7.3</td>
</tr>
<tr>
<td>Ramp Area</td>
<td>7.2</td>
</tr>
<tr>
<td>Runway</td>
<td>7.0</td>
</tr>
<tr>
<td>Airline</td>
<td>7.1</td>
</tr>
</tbody>
</table>
DEPARTURE TAXI TIME PREDICTIONS
USING ASDE-X SURVEILLANCE DATA

<table>
<thead>
<tr>
<th>Meteorological Cond.</th>
<th>7.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDQ</td>
<td>5.1*</td>
</tr>
<tr>
<td>Predicted VDQ</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Our contention that the situational awareness provided by ASDE-X creates the potential for improvement in taxi-out predictions is supported by the data; however, much work remains to be done.

First, algorithms to better predict the actual VDQ should be developed. These algorithms will need to account for the confluence points on the airport surface where aircraft may be dynamically re-ordered by ATC.

Second, the variations in taxi-out time still present in Figure 6 indicate that variability in inter-departure time must be accounted for. This variability may be partially explained by wake-vortex weight classes and miles-in-trail restrictions in the Terminal Radar Approach Control (TRACON) airspace.

Third, it is possible that some taxi-out variation is simply an artifact of the measurement system. In particular, it would be helpful to better understand the precise mechanism(s) responsible for the BTS out-time.

Fourth, this analysis should be ported to another airport, preferably one with better ramp area and gate surveillance.

Fifth and finally, refinement of the data repository this work is based upon – in particular, an investigation into the potentially missing operations indicated by ASPM and/or BTS data.

Acknowledgements

We gratefully acknowledge the support, direction, and review of our work that had been generously provided by our colleagues at Sensis Corporation. In specific, we thank Mark Runnels, Collen Knickerbocker, and Nichole Wenderlich.

References


Copyright Statement

The authors confirm that they, and/or their company or institution, hold copyright on all of the original material included in their paper. They also confirm they have obtained permission, from the copyright holder of any third party material included in their paper, to publish it as part of their paper. The authors grant full permission for the publication and distribution of their paper as part of the ICAS2008 proceedings or as individual off-prints from the proceedings.