A STATISTICAL LEARNING APPROACH TO THE MODELING OF AIRCRAFT TAXI TIME

Richard Jordan, Mariya A. Ishutkina, Tom G. Reynolds,
MIT Lincoln Laboratory, Lexington, MA

Abstract

Modeling aircraft taxi operations is an important element in understanding current airport performance and where opportunities may lie for improvements. A statistical learning approach to modeling aircraft taxi time is presented in this paper. This approach allows efficient identification of relatively simple and easily interpretable models of aircraft taxi time, which are shown to yield remarkably accurate predictions when tested on actual data.

1. Introduction

An air traffic control tower advanced automation system known as the Tower Flight Data Manager (TFDM) is being considered for development by the Federal Aviation Administration (FAA) [1]. TFDM is designed to replace the numerous standalone systems within current air traffic control towers with an integrated technology suite combining new surveillance displays, flight information management and decision support tools. These integrated technologies offer the potential to enable multiple system benefits, including reduced delay, taxi time and fuel burn (with associated economic and environmental impacts). But in order for the system to deliver these benefits, its developers need to understand key airport operating characteristics and then transform this knowledge into value-added TFDM functionality. Airport taxi processes are important in this regard because delay and fuel burn often manifest during taxi in the current system and hence is where benefits may be delivered with new systems.

The development of accurate and robust algorithms for taxi time prediction has been identified as a TFDM priority. Such algorithms are envisioned to play a critical role in, among other things, the optimization of departure sequences with respect to delay and throughput, the timing of airport configuration changes, and the efficient and timely delivery of departing aircraft to the runway in order to take advantage of route openings in the overhead stream. The latter is especially pertinent during inclement weather and/or heavy traffic volume, where the time window during which an aircraft must take off and merge into the overhead stream can be very narrow.

Recent publications have studied taxi processes and the prediction of taxi times in some detail [e.g. 2, 3, 4]. This paper presents a new approach to taxi process modeling specifically intended to support the needs of the TFDM development activity. A statistical learning method is followed to extract key predictor variables from a set of many. The variable selection method automatically performs a trade-off between the number of variables included in and the performance of the taxi time models. The resulting models not only provide accurate predictions, but they also help us identify and focus on key variables that impact an aircraft's taxi time.

The paper is organized as follows. Section 2 provides a description of the key characteristics of Dallas/Fort Worth Airport (DFW) (the focus for model development because it is the location for TFDM prototype efforts) and the data available at that site. In Section 3, a standard linear regression model for taxi-out processes based purely on taxi distance at DFW is presented to act as a baseline for comparison with the new model. In Section 4, other candidate variables (in addition to distance) are explored. These are then utilized by the statistical learning approach for taxi time modeling, which is described in Section 5, together with results of its application to taxi-out processes at DFW. In Section 6, we use the same learning methodology to develop a model for taxi-in time. Finally, conclusions and directions for future work are presented in Section 7.

2. DFW Airport Characteristics

The models described in this paper were developed for Dallas/Fort Worth International Airport. The airport layout is shown schematically in Figure 1.
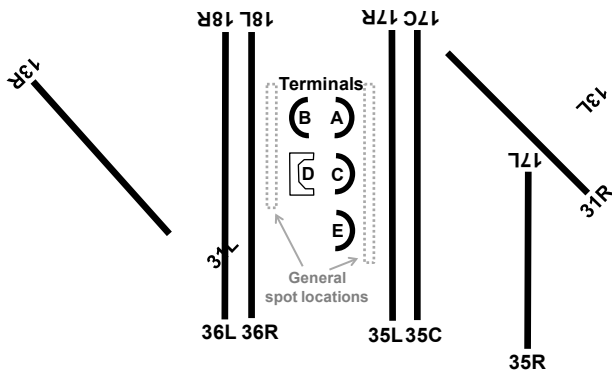
Figure 1. Dallas/Fort Worth Airport Schematic

The airport handles about 1,800 operations per day. The majority of the flights land or depart from the four parallel runways located around the central terminal area. The inner runways (17R/35L and 18L/36R) are typically used for departures, while the outer runways (17C/35C and 18R/36L) are typically used for arrivals. In this paper, the analysis covers taxi-out to runway 17R and taxi-in from 17C. Figure 2 shows the aerial view of the central terminals and the four parallel runways. The topology of the airport is such that arrivals landing on 17C have to cross 17R while the departures taxiing from the central terminals to 17R have no runway crossings.
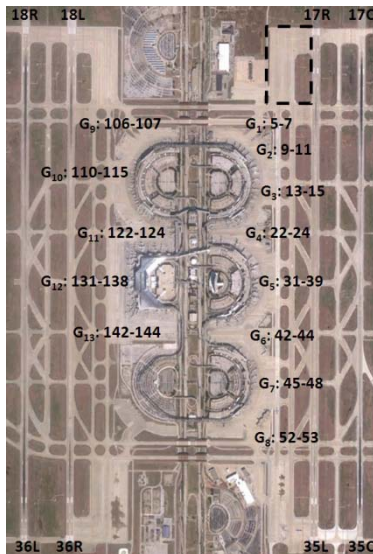

Figure 2. Aerial View of DFW[1] Showing Runways, 17R Queue Box and Spot Groupings

The departure traffic at DFW is controlled by the Ramp Controller in the ramp area, by the Ground Controller from the spot to the runway queue and by the Local Controller from the queue to wheels-off. Analogously, the arrival traffic is controlled by the Local Controller on the runway through the high-speed runway exit, the Ground Controller from the high-speed exit to the ramp area, and the Ramp Controller to the gate. The taxi-out model presented below predicts the taxi time from spot to the runway queue, while the taxi-in model predicts the taxi time for arrivals from the arrival runway to the spot. In addition, we are in the process of developing taxi-out models from spot all the way to wheels-off. The results of this research will be presented elsewhere.

We used archived data from the Runway Status Lights (RWSL) system for analysis of taxi operations at DFW. This data set consists of surveillance data for aircraft and vehicles on or near the airport surface detected by the Airport Surface Detection Equipment Model–X (ASDE-X) system. ASDE-X is a surface surveillance system that fuses the data from multiple surface movement radars, transponder multilateration sensors and Automatic Dependent Surveillance-Broadcast (ADS-B) sensors at 1Hz update rate and approximately 8m lateral resolution [5]. ASDE-X does not classify tracks as ground vehicles, arriving or departing aircraft.

Analysis of tracks was conducted by matching the time a given aircraft or vehicle spent while traversing or waiting at a particular airport feature. The following features were used: spot boxes in the ramp area, static runway queue boxes and runway thresholds.

Spot boxes were defined around the official spot locations. The dimensions of the spot boxes were chosen to accommodate the variability of transponder placement between aircraft (at least twice the typical length of a heavy aircraft) and to completely cover the ramp exit area from each terminal. This enables our data processing algorithms to detect aircraft that move to the active airport surface without passing through an official spot location. Given that there are more than 50 spot locations at DFW, for the purposes of taxi-in modeling described in Section 6, the spots are grouped according to proximity as shown in Figure 2.

---

[1] Source: satellite image from GoogleEarth.

We have defined queue boxes in such a way as to encompass the area around each of the four main runways while excluding any taxi ways. For example, the queue box for 17R is shown in Figure 2; its dimensions are 2000 x 1000 ft. Queue waiting times and runway occupancy times were also analyzed, but the results of that analysis are not included in this paper.

The runway threshold time for arrivals is defined as the first time an aircraft is detected within a horizontal boundary of a particular runway. Arrival threshold time is used as a proxy for wheels-on time because altitude data was found to be especially noisy for aircraft during the landing process.

Our processed ASDE-X DFW dataset contains approximately 900 departures and arrivals per day, which agrees well with the numbers in the Aviation System Performance Metrics (ASPM) database[2]. But the ASDE-X data has some limitations. First, tracks may be missing data. In particular, because the pilots are required to turn on transponders only once they are on the active airport surface, ramp area coverage is often lacking. Overall, almost a third of departures and a quarter of arrivals have missing spot information and hence have to be excluded from the analysis of taxi times. For taxi-out analysis we used only flights that were detected leaving a spot and entering a particular queue box. Second, it is sometimes difficult to separate fused tracks when several arrivals and departures share a given call sign. Third, it can be difficult to differentiate ground vehicles from aircraft, as tracks may be missing altitude information. To circumvent these data limitations, tracks for flights that resulted in taxi times that differed by more than three standard deviations from the mean for given taxi distances and taxi routes were checked to ensure that only the actual aircraft data are included in our model development and analysis.

The following Visual Meteorological Conditions (VMC) mid-week days were used as training days for the models: 11/02/09 through 11/05/09 plus 09/01/09. Taxi-out and taxi-in models were developed for runways 17R and 17C, respectively. In total, 3166 departure flight tracks and 1554 arrival flight tracks were included in the training dataset. The following

VMC mid-week days were used for testing the models: 07/13/09 and 11/27/09. For the former day, the data includes 341 departures and 262 arrivals, while for the latter there are 299 departures and 226 arrivals. We note during all the test days and training days, the airport operated almost exclusively in southflow. In a typical southflow operation runways 18L, 17R, 13L are used for departures, while runways 18R, 13R, 17C and 17L are used for arrivals.

3. Standard Linear Regression Model

We first present the results of a standard ordinary least squares (OLS) linear regression model for the spot→queue time, T_{out}, in which the distance, D, traveled from the spot to the queue box is the single independent variable. While we discuss the results only for runway 17R at DFW, we've found that the regression equation for runway 18L is essentially identical to that for runway 17R (the slopes of the regression lines agree to two decimal places and the intercepts differ by less than 9%). In addition, there is no significant difference in the level of prediction accuracy, as measured by the usual statistics, such as mean absolute error, mean squared error or R^2.

The regression equation for spot→queue time for runway 17R is:

$$T_{out, ols} = 1.925D + 0.367 \qquad (1)$$

Here, D is measured in kilometers and $T_{out, ols}$ is in minutes. Figure 3 shows this regression line superimposed on a scatter plot of observed (D, T_{out}) pairs for runway 17R on the independent test day 07/13/09. We see that 85.3% of the predictions based on this simple model are accurate to 1 minute and 97.9% are accurate to 2 minutes. The model yields very similar results for the test day 11/27/09 (see Table 3 below).

This analysis demonstrates that distance alone is a reasonably strong predictor of the taxi time from spot to queue. In fact, the correlation between the taxi time and taxi distance in the training data is 0.89. Nevertheless, a close examination of Figure 3 reveals that, while the model is quite accurate for short and intermediate taxi distances, it cannot capture the spread in observed taxi times for taxi distances of approximately 2 kilometers and beyond. The

---

[2] For detailed cross-validation analysis of ASDE-X and ASPM data, see [5].

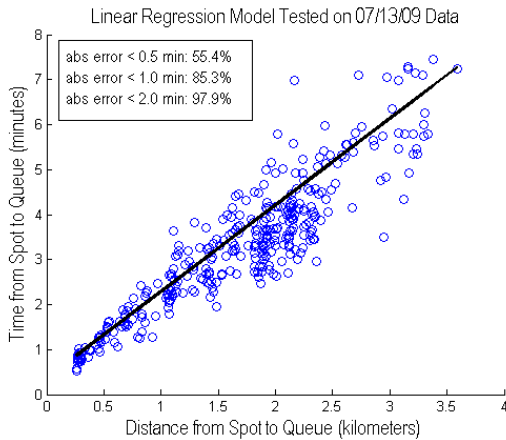implication is that factors other than distance are at play.

Figure 3. Standard Linear Regression Model Tested on 07/13/2009

In the next sections, a modeling approach is described that aims to build more accurate models via the extraction of the most significant variables from a list of many.

4. Independent Variables and Basis Functions for Taxi-Out Modeling

Based on the results of the previous section, we wish to identify factors responsible for the variability in taxi-out times evidenced by Figure 3. For spot→queue taxi time prediction, our goal is to build models based on variables that will either be known exactly at the time an aircraft is at its spot or else could be estimated with high accuracy via surveillance, flight schedules, etc. The challenge is in deciding which, of many, variables to include. Our approach is to add to the distance variable a relatively large set of additional variables that may impact an aircraft's taxi out time and then apply a method that builds up the regression equation by recursively including and excluding basis functions according to whether they improve the model's predictive ability. Basis functions are candidate independent variables and whatever functions—e.g. products, powers—of them are to be considered for inclusion in the regression equation. Such approaches, referred to as *subset selection*, are commonly used in machine learning and pattern recognition for feature selection/extraction [6]; their application to statistical learning and regression analysis is a burgeoning field of research, where they have been shown to produce accurate and easily interpretable models involving only a small number of basis functions [7, 8, 9]. We follow such an approach largely because we find it difficult to know *a priori* which variables are likely to have the most significant impact on taxi-out time. The method we use will be described in the next section.

First we must decide on a list of independent variables. There are many variables that could conceivably have an impact on taxi processes, and the objective here is to identify a pool of initial candidate variables (e.g. by examination of Figure 3 and ASDE-X track data). One possible explanation for the spread of spot→queue taxi times for a given distance (as seen in Figure 3) is a difference in typical taxi speeds for different airlines. Such differences would be more pronounced over long taxi distances. Accordingly, we analyzed the average taxi speeds for aircraft that, based on the ASDE-X tracks, traveled from various spots to runway 17R. Table 1 shows the results for two different airlines that taxied from spot 142 on the western side of the airport on 11/02/09—11/05/09.

Table 1. Comparison of Average Taxi Speeds

Airline	Avg Taxi Speed from Spot 142 to Queue at 17R	Standard Deviation	# of Flights
A_1	16.1 knots	1.35 knots	10
A_8	19.0 knots	1.34 knots	18

Figure 4 demonstrates that all the aircraft in question followed nearly identical taxi routes and, therefore, taxied roughly the same distance. As this provides strong evidence that there are considerable differences in airline taxi-out speeds, even for a given taxi route and distance, we include in our list of independent variables a binary variable associated with each airline for which there are an adequate number of data points in the training dataset. Specifically, we include variables $A_1, \ldots, A_{13}$, where $A_j = 1$ for airline carrier j and $A_j = 0$ otherwise. We note that the airline variable may be a proxy for aircraft type and this will be explored further in the future.
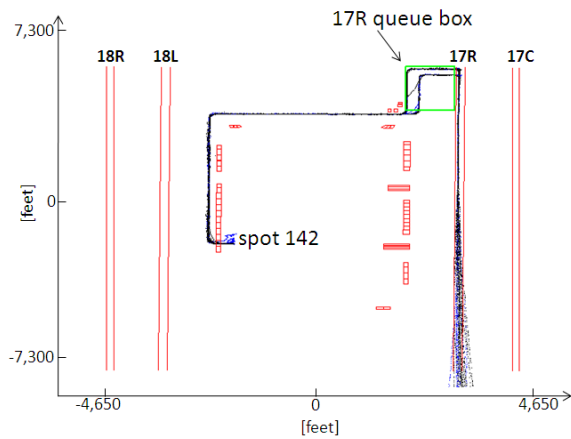
Figure 4. Taxi Routes from Spot 142 to 17R

In addition, analysis of the ASDE-X data reveals that, for a given taxi distance range, aircraft that taxied from the west side of the airport tend to take longer than those that taxied from the east side. For example the average taxi time for aircraft that taxied a distance between 2.5 km and 3.0 kilometers to runway 17R from the west side of the airport is 5.89 minutes, while the average for aircraft with taxi distances in this range but taxiing from the east side of the airport is 5.26 minutes. While part of this discrepancy may be due to differences in taxi speeds for different airlines, we include in our set of candidate variables a binary variable W, which is equal to 1 if the aircraft taxies from the west side and 0 if it taxies from the east.

At least intuitively, surface congestion encountered while taxiing from spot to queue is also expected to have an impact on an aircraft's taxi time. Therefore, we include two congestion variables in the list of independent variables. These are N_{sd} and N_{ath}, which are defined as follows:

- N_{sd} (i) is the number of departing aircraft that leave a spot and taxi to runway 17R during the time interval $[s(i) – 1.67, s(i) + 0.50]$ minutes, where $s(i)$ is the time that aircraft i left its spot to taxi to runway 17R.
- N_{ath}(i) is the number of arrival aircraft that landed (or, more exactly, crossed the arrival runway threshold) on the east side of the airport during the interval $[s(i) – 1.58, s(i) + 0.75]$ minutes, where $s(i)$ is as above.

The time interval in the definition of N_{sd} has been chosen to (approximately) maximize the correlation between T_{out} and N_{sd}. Precisely, for each aircraft i, we determined via analysis of the ASDE-X data the number of departures $M(i, j, k)$ that left their spot and taxied to runway 17R during the time interval $[s(i) - j/12, s(i) + k/12]$, for $j, k = 0, 1, …,180$ (so the maximum "before" and "after" times are 15 minutes) and determined k and j such that the vector $M(:, j, k)$ is maximally correlated with the vector T_{out}. N_{sd} is defined to be the vector that achieves this maximum correlation. Based on the data in the training set, the maximum correlation is 0.24, that is, $corr(T_{out}, N_{sd}) = 0.24$. The time interval in the definition of N_{ath} is found by the same procedure, and $corr(T_{out}, N_{ath}) = 0.11$. Note that these are small correlations, which suggests that surface congestion, at least measured the way we have done so, may not have played a major role in the spot→queue times for runway 17R on the days represented in the training set. Nonetheless, we include N_{sd} and N_{ath} in our list of independent variables.

Hence, the list of independent variables is:

- D
- $A_1,…,A_{13}$
- W
- N_{sd}
- N_{ath}

As the models are further developed and applied to different situations (e.g. at DFW and at different airports), the pool of candidate variables is likely to be refined. But given this list of independent variables identified above, we can define the set of basis functions for consideration in the regression model. Because we believe that interactions between the binary variables and the other variables may be important (for example, a term in the regression equation such as xA_mD will imply that x minutes should be added or subtracted, depending on the sign of x, per kilometer distance to the taxi out time for airline m) we take as the set of basis functions 1, D, $A_1,…$, A_{13}, W, N_{sd}, N_{ath} and quadratic interactions between them. We do not consider higher order interactions, as we wish to avoid complexity and over-fitting in the model. Note that the airline variables are orthogonal, so that any product between them is identically 0, and is, therefore, not part of the

1.B.1-5

basis set. Also, as $A_1, \ldots, A_{13}$, and W are binary variables, so they are equal to their squares. Thus, the total number of possible basis functions is 79.

5. Sequential Forward Floating Subset Selection (SFFSS) Method for Taxi-Out Modeling

One way to proceed in developing a taxi-out model would be to perform an exhaustive search over the set of all the basis functions described above to determine the linear combination that achieves the maximum of some "goodness of fit" objective function J on the training data. However, given the dimension of the basis set, such an approach is, for all practical purposes, infeasible. Instead, we use the Sequential Forward Floating Subset Selection (SFFSS) method [7] to extract from the large set of basis functions a small subset that captures much of the variability in the data and yields small prediction errors. Here is a basic description of the SFFSS algorithm.

Suppose that the current model includes and excludes a certain number of basis functions:

- **Step 1: Inclusion.** Select from the list of remaining basis functions the one whose inclusion leads to the greatest increase of the objective function J. Add it to the existing list of included basis functions.

- **Step 2: Conditional exclusion**. Find the least significant basis function among those included thus far. If it is the basis function just added, then keep it and return to Step 1. Otherwise, exclude it and continue to Step 3.

- **Step 3: Continuation of conditional exclusion**. Again find the least significant basis function among those included thus far. If its removal will (a) leave at least two basis functions, and (b) the value of J is greater than that for any subset of basis functions of the same size already found, then remove it and repeat step 3. When these two conditions cease to be satisfied, return to step 1.

The search terminates when no further increase in J is achieved. It starts with the constant 1 as the only basis function and proceeds from there. By "least significant basis function", we mean the one whose removal least decreases the value of J. The inclusion and exclusion search methods are forward selection and backward selection, respectively. A detailed description of these methods may be found in [10].

The objective function J we use is $-n\log(\text{GCV})$, where GCV is the generalized cross-validation, defined as GCV $= \text{MSE}/(1 - m/n)^2$. That is,

$$J = 2n\log(1 - m/n) - n\log(\text{MSE}) \qquad (2)$$

Here n is the dimension of the entire basis function set, m is the number of basis functions currently under consideration, and MSE is the mean squared error associated with the given model. This objective function provides a judicious trade-off between dimensionality and the model's predictive capacity. GCV was introduced in [11, 12] and is commonly used in practice. Our transformation and scaling of it is for computational efficiency and stability purposes.

We remark that SFFSS is a sub-optimal method. There is no guarantee that it finds the global optimum. The method, however, has been shown to yield accurate results and to outperform many other sub-optimal approaches [7].

Applying the SFFSS approach to the training data, we obtain the following regression equation for taxi-out time from spot to queue for runway 17R at DFW:

$$T_{\text{out, sffss}} = 0.290 + 1.832D - 0.684A_6 + 0.436W + 0.146A_1D - 0.256A_7D - 0.336A_8D + 0.043N_{sd} \qquad (3)$$

This equation contains only 8 terms (equal to the number of surviving basis functions) and makes use of only 7 out of the 18 possible independent variables. As a result, it is easy to interpret. For example, the 5th term on the right hand side indicates that 0.146 minutes per kilometer should be added for airline A_1, while the 7th term says that 0.336 minutes per kilometer should be subtracted for airline A_8. We note that these results are consistent with the average speeds in Table 1.

While the congestion variable N_{sd} is included in the equation, it is the variable of least impact. The maximum N_{sd} observed in the training data and the test data is 8, which amounts to an additional taxi

time of only about 20 seconds. This would be significant for short taxi distances, but not for longer taxi distances.

We provide an example prediction using Eqn. (3). Suppose airline A_1 taxies from spot 132 to runway 17R. From analysis of the ASDE-X data, the average taxi distance in this case is approximately 2.6 kilometers. We have $D = 2.6$, $W = 1$ and $A_1 = 1$ and all the other variables are 0, except, perhaps, N_{sd}. With $N_{sd} = 0$, we obtain from (Eqn. 3) the predicted taxi time of 5.87 minutes. With $N_{sd} = 7$, the maximum observed value on 07/13/09, we obtain a predicted taxi time of 6.17 minutes. The average taxi time for airline A_1 from spot 132 to runway 17R on this day was 5.95 minutes. Thus the model prediction is within 0.22 minutes (or 13.2 seconds) of the average time. More importantly, the model significantly outperforms the mean taxi-out time as a predictor over the range of distances represented in the data.

Tables 2 and 3 compare the predictions of this model with those of the linear regression model (Eqn. 1) for the test days of 07/13/09 and 11/27/09, respectively.

Table 2. SFFSS vs. OLS for 07/13/09

Model	Abs error < 0.5 min	Abs error < 1.0 min	Abs error < 1.5 min	Abs error < 2.0 min	R^2	Mean abs error	Max abs error
SFFSS	65.7%	91.2%	97.4%	99.4%	0.92	0.44 min	2.41 min
OLS	55.4%	85.3%	96.5%	97.9%	0.82	0.53 min	2.80 min

Table 3. SFFSS vs. OLS for 11/27/09

Model	Abs error < 0.5 min	Abs error < 1.0 min	Abs error < 1.5 min	Abs error < 2.0 min	R^2	Mean abs error	Max abs error
SFFSS	69.0%	90.6%	97.3%	100%	0.93	0.43 min	1.77 min
OLS	59.8%	83.7%	96.7%	97.9%	0.79	0.55 min	2.59 min

Clearly the SFFSS model (Eqn. 3) gives more accurate predictions than OLS (Eqn. 1). The most improvement is in the percentage of predictions accurate to 0.5 minutes (improvement of 10.3% for 07/13/09 and 9.2% for 11/27/09) and 1.0 minutes (improvement of 7.3% for 07/13/09 and 6.9% for 11/27/09). In addition, SFFSS produces larger R^2

and smaller mean and maximum absolute error for both test days.

Figure 5 shows the performance of the SFFSS model as a function of taxi distance for 07/13/2009 data. For comparison, the linear fit to distance according to the OLS model is also included in the graph.
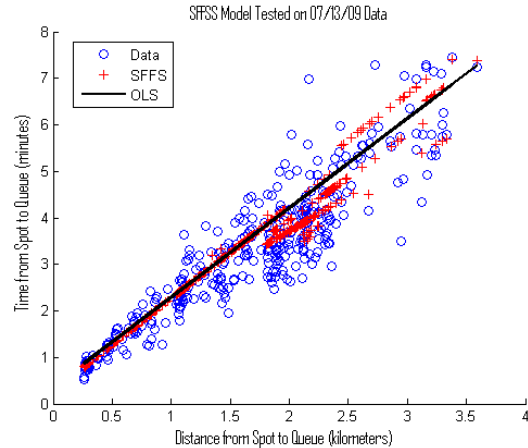

Figure 5. SFFSS Model Tested on 07/13/09

Observe that the SFFSS model splits into multiple "branches", which is especially obvious for distances on the order of 2 kilometers and greater. This branching allows it to more accurately capture taxi times above and below the overall mean for a given distance, and this is what accounts most for its increased accuracy over OLS.

6. SFFSS Method for Taxi-in Modeling

We have also used the SFFSS approach to construct a model of taxi-in time for runway 17C. By taxi-in time, we mean the time between when the aircraft crosses the arrival runway threshold (a proxy for wheels-on) and the time it arrives at its spot. Because most aircraft that land on runway 17C cross the departure runway 17R, the modeling of taxi-in time from runway 17C represents a greater challenge than modeling the spot→queue time for departures on runway 17R. In particular, the interplay between departures and arrivals is expected to be important, insofar as arrivals may stop and wait at the crossing for departures occupying the runway.

In modeling the taxi-in time, we include among the list of independent variables the binary variables for each of the 13 airlines considered in the taxi-out model. Rather than taxi distance, we use (runway,

spot grouping) pairings as a proxy. This would be an inadequate proxy if multiple taxi routes with widely varying distances are used from runway 17C to a given spot. But our analysis of the ASDE-X data indicates that this is not the case at DFW. We use spot groupings rather than distance in the construction of our taxi-in model in part out of curiosity. We wish to gauge the predictive capacity of such a model. We also have in mind the fact that taxi distance can only be approximated in real time, while the spot, or at least the terminal should be known exactly. The variables $G_1,\ldots,G_{13}$ are binary variables for the spot groups which are shown in Figure 2.

We also include two congestion variables. N_{dep} is equal to the number of departure aircraft that leave their spot and taxi to a runway on the east side of the airport up to 5.7 minutes before or 3.3 minutes after the arrival aircraft in question crosses the threshold at runway 17C and N_{arr} is the number of arrival aircraft that cross the threshold at runway 17C or runway 17L up to 1.3 minutes before the aircraft in question. The time intervals in the definitions of these variables have been chosen so as to maximize the correlation with the taxi-in time T_{in}, using the same method that was used to define the congestion variables for taxi-out time modeling, as described above.

Hence, the list of independent variables for the taxi-in model are:

- $A_1,\ldots,A_{13}$
- $G_1,\ldots,G_{13}$
- N_{dep}
- N_{arr}

As before, the set of basis functions consists of the constant 1, these variables and quadratic interactions between them. Applying the SFFSS method, we obtain the following regression equation for the taxi-in time from runway 17C to spot:

$$T_{in,\,sffss} = 6.751 - 0.819G_4 - 1.652G_5 - 2.970G_6 - 3.158G_7 - 2.020G_8 + 2.099G_{10} + 1.330G_{11} + 3.711G_{12} - 0.973A_1 - 0.789A_4 - 1.076A_5 + 0.447A_9 + 2.687A_1G_{11} - 0.478A_7G_7 + 2.159A_1G_{13} - 1.052A_{11}G_7 + 0.092N_{dep} + 0.296N_{arr} \qquad (4)$$

This equation makes use of 17 out of the 28 possible independent variables and 19 basis functions. Note that the congestion variables play a more important role in this model than they did in the taxi-out model, as expected given the greater interactions occurring for arrivals. For example, in the training data N_{dep} is seen to be as large as 12. This amounts to an additional 1.10 minutes in taxi-in time, according to Eqn. (4). In general, the 18th term on the right hand side of the equation represents the expected additional taxi-in time for an arrival aircraft that encounters departure congestion of the magnitude N_{dep}. It does not tell us where during taxi-in this delay is expected to be incurred—i.e., whether or not it is due to the aircraft stopping at the crossing to wait for departures occupying runway 17R. A more sophisticated model would be needed for that purpose. Similarly, the 19th term on the right hand side of Eqn. (4) represents the excess taxi-in time due to arrivals that landed on the east side of the airport up to 1.3 minutes prior to the aircraft in question. In the training data, this number is no larger than 3, which amounts to a predicted additional taxi-in time of 0.89 minutes.

Table 4 summarizes the predictive accuracy of this model based on data from 07/13/09. Essentially identical results are obtained for the 11/27/09 data.

Table 4. SFFSS Taxi-In Model Tested on 07/13/09

Abs error < 1.0 min	Abs error < 1.5 min	Abs error < 2.0 min	Abs error < 3.0 min	R^2	Mean abs error	Max abs error
66.3%	83.4%	90.9%	98.3%	0.78	0.89 min	4.54 min

The accuracy of the taxi-in model is not as high as that for the spot→queue taxi-out model. For example, the coefficient of determination is $R^2 = 0.78$ for the model tested on 07/13/09 data, whereas R^2 for the taxi-out model exceeds 0.9 for both test days. This decreased accuracy is to be expected given the interaction of arrivals at the crossings at runway 17R with departures using the runway. It is possible that a model that uses taxi-in distance rather than (runway, spot group) pairings would be more accurate, but we believe that a critical next step is to capture more intelligently the interactions of arrivals with departures at the runway crossing.

The present model, however, does outperform the mean taxi-in time for each spot grouping as a predictor, at least on the test days of 07/13/09 and 11/27/09. To illustrate, on the latter of these two days, the ASDE-X data shows that 30 aircraft landed on runway 17C and taxied to the spot group G_7. The

average taxi-in time for these aircraft was 3.56 minutes and 19 of them, or 63.3% had taxi-in times within 1 minute of the mean. Equation (4) provides predictions of taxi-in times within 1 minute accuracy for 23 of these aircraft, or 76.7%. More generally, we believe that the results demonstrated in Table 4 show that the SFFSS approach holds promise for taxi-in modeling and warrants further exploration.

7. Conclusions and Future Work

The results presented in this paper show that statistical learning techniques can be used to produce accurate, yet relatively simple models of aircraft taxi time. The models presented above were developed for Dallas/Fort Worth airport on good weather days. These models or refinements thereof will be used for real-time taxi time predictions to support TFDM decision support tool development.

In the nearest future we plan to incorporate other runways, other weather conditions and queue waiting times for our current DFW application. We will also extend this model to accommodate differing topologies and traffic levels at other airports. As part of this work, the impact of other independent variables will be evaluated, including time of day, day of week, aircraft type, weather/visibility conditions, and number of runway crossings per taxi route. SFFSS and other subset selection methods [9] are expected to be especially valuable in this process, given their propensity to efficiently extract a small set of highly significant variables from a much larger list of candidate variables.

Finally, we point out that while SFFSS models presented above provide accurate predictions on the test data, it will be necessary to use multiple approaches and compare their predictive accuracies when additional independent variables are incorporated. A close competitor to SFFSS for the particular training and test data and the independent variables we have considered thus far is the M5P model tree approach [12, 13 and 14]. The results of our models developed using this approach will be presented in a subsequent manuscript.

References

[1] Nene, V. D. and C. E. Morgan, 2009, A Mid-Term Concept of Operations for a Tower Flight Data Manager (TFDM), MITRE MP090169.

[2] Gong, C., 2009, Kinematic Airport Surface Trajectory Model Development, AIAA 2009-7076, 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO).

[3] Balakrishna, P., R. Ganesan and L. Sherry, 2009, Application of Reinforcement Learning Algorithms for Predicting Taxi-out Times, Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM2009).

[4] Kistler, M. S. and G. Gupta, 2009, Relationship Between Airport Efficiency and Surface Traffic, AIAA 2009-7078, 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO).

[5] Legge, J., B. Levy, 2008, Departure Taxi Time Predictions Using ASDE-X Surveillance Data, 26th International Congress of the Aeronautical Sciences.

[6] Guyon, I. and A. Elisseeff, 2003, An Introduction to Variable and Feature Selection, Journal of Machine Learning Research 3, pp. 1157-1182.

[7] Pudil, P., K. Fuka, K. Beránek, P. Dvorák, 1999, Potential of Artificial Intelligence Based Feature Selection Methods in Regression Models, Third International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'99), pp. 159-168.

[8] Hastie, T., R. Tinshirani and J. Friedman, 2009, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Edition, New York, Springer.

[9] Yan, X. and X. Gang Su, 2009, Linear Regression Analysis: Theory and Computing, Singapore, World Scientific.

[10] Dash, M., and H. Liu, 1997, Feature selection for classification, Intelligent Data Analysis – An International Journal, Volume 1, Number 3, pp, 131–156.

[11] Craven, P. and G. Wahba, 1979, Smoothing Noisy Data with Spline Functions, Numerische Mathematik, 31, pp, 377-403.

[12] Golub, G. H., M. Heath, and G. Wahba, 1979, Generalized Cross-Validation as a method for choosing a good ridge parameter, Technometrics, 21(2), pp. 215-223.

[13] Wang, Y. and I. H. Witten, 1996, Induction of Model Trees for Predicting Continuous Classes, Working Paper 96/23, Working Paper Series, Department of Computer Science, University of Waikato, New Zealand.

[14] Quinlan, J. R., 1992, Learning with Continuous Classes, Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Singapore, World Scientific, pp. 343-348.

Acknowledgements

Disclaimer

Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

29th Digital Avionics Systems Conference
October 3-7, 2010