9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)
and
Air
21 - 23 September 2009, Hilton Head, South Carolina

AIAA 2009-7078

Relationship between Airport Efficiency and Surface Traffic

Matthew S. Kistler[*] *and* Gautam Gupta[†]
University of California, Santa Cruz, NASA Ames Research Center, Moffett Field, CA, 94035

The focus of this paper is to identify and validate relationships between an airport efficiency metric and aggregate factors related to surface traffic movement. For validation, data from Dallas/Fort Worth (DFW) International airport is analyzed. Taxi time is used as the metric of efficiency with aggregate surface traffic count, taxi distance, and number of stops identified as large contributing factors to inefficiency. Simple linear and log-linear functional forms are used in regression to find the effect these factors have on taxi time, with variations of both models fitting the data with adjusted R^2 values greater than 0.95. Predictive capability of the models was tested on an independent dataset. Linear models estimated 71% of the taxi times within one minute of the observed data while log-linear models estimated just fewer than 65% of the taxi times within one minute of the observed data. Estimates and prediction results indicate the need for testing alternate functional forms for the relationship between taxi time and the above mentioned factors of efficiency.

I. Introduction

With the predicted increase in air traffic in the future[1], there is a need to use the existing infrastructure in an efficient manner. Efficient usage is especially critical for physical infrastructure assets like runways and taxiways, since expanding these poses financial, operational and environmental challenges. To develop tools assisting future airport operations under increased demand, metrics of airport efficiency need to be identified (both operational efficiency as well as environmental efficiency).

Various taxi phenomenon can be considered as metrics of efficiency, including aircraft stops or stopping time on surface. Taxi time can be one such measure of efficiency and would include stops itself. An aircraft's taxi time is governed by minute surface traffic details, including the number of "conflicting" aircraft passed, the number of aircraft already present on the route taken and others. Previous work on studying microscopic surface traffic data has focused on either studying the existing sequencing strategies[2] or identifying and modeling the stochastic nature of taxi trajectories[3]. A previous study on the benefit of multiple surface initiatives using traffic demand depicts the effect of the number of taxiing aircraft on mean taxi time and throughput[4]. This model shows the benefit of adding ASDE-X equipment and a new runway to Orlando International Airport (MCO). A major drawback of that model is the consideration of only the surface traffic demand as the relevant factor, and the use of mean taxi times, instead of actual taxi times used in the analysis. Another study proposes and tests methods for filtering surveillance data for better estimates of aircraft states while taxiing[5].

This paper uses taxi time as one such metric for airport efficiency. Further, even though taxi time would depend on traffic details, it can potentially be modeled as a function of *aggregate* surface traffic parameters (for example, number of aircraft taxiing, arrival and departures rates and others), circumventing the need for minute taxi-trajectory analysis. Studying such models of taxi time has two major benefits: First, the marginal or additional contribution of each factor towards increased taxi time (inefficiency) is identified, thus highlighting the sensitivity of each factor. This can be seen as a long-term benefit, where the relationship is used in planning taxi trajectories. Second, such aggregate factors (like the total traffic count on the surface or the count in a particular section) are direct observables for the air traffic controllers. If an observable is not favorable, controllers could limit the access to the taxiways till the observable is within the efficiency zone, thus providing guidelines to improving efficiency of surface operations by limiting or changing some procedures. Based on these, various aggregate factors affecting taxi time are identified in this paper. The factors identified include aggregate traffic counts as well as the aircraft's surface route information, including number of turns and distance traveled. The relationship between the factors and taxi time was then tested using microscopic surface traffic data, which includes flight data on the airport surface excluding the

[*] Aerospace Engineer, University of California, Santa Cruz, Moffett Field, CA, USA
[†] Associate Research Scientist, University of California, Santa Cruz, Moffett Field, CA, USA

terminal area. Rather than predicting taxi-times itself, the aim is to model relationships between a microscopic efficiency metric (taxi time) and macroscopic traffic factors.

The rest of the document is organized as follows: the next section describes the data source and methodology for collecting the data for taxi time and the influencing factors. This is followed by analysis of the relationship between metric and factors. The analysis is divided into three parts: first the "spread" of certain data is characterized. Then the results from testing different models of the relationship are presented. In the third subsection, the estimated models are tested over an independent dataset. The paper concludes with directions for future work.

II. Methodology

As discussed before, taxi time is chosen as a metric of efficiency as various cost (fuel savings) and environmental (exhaust) effects can be directly attributed to taxi times. It is possible to choose other metrics, for example the time an aircraft stops during taxiing or the number of stops itself. However, it is difficult to select a functional form for these metrics and express efficiency in a quantifiable manner. The time to move from point A to B can be based on kinematics with some allowance for congestion effects. But stopping time and number of stops are highly dependent on situational inputs, leading to large errors in estimation. Preliminary tests on using these are measures of efficiency showed no significant relationship with aggregate surface parameters. Thus, taxi time is used as a measure of efficiency.

There are many factors that can influence the taxi time of an aircraft. The factors chosen in the analysis are values that can be placed in aggregate form, and can be observed by a controller from an air traffic control tower. These factors include aggregate surface traffic count, flow rates (arrivals and departures) and trajectory geometry (taxi distance, number of turns). In this section the data collection effort is described first, including the layout and the method for calculating the relevant quantities. Then the models used in the analysis are described.

A. Airport Layout and Data Collection

Dallas/Fort Worth (DFW) International airport was used as a case study for this analysis. It was split into four regions, as shown below in Figure 1, for the purpose of narrowing down the effect of surface traffic location on taxi time.

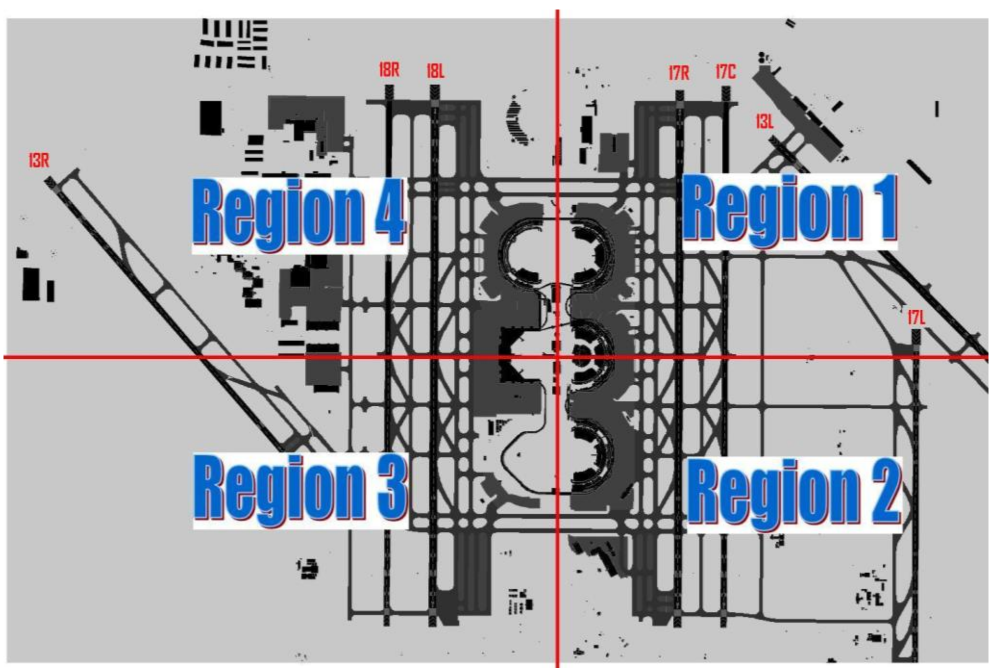

Figure 1. Dallas/Fort Worth (DFW) International Airport Layout and Regions for Location Effects

Surface data at DFW airport was collected using the Surface Operations Data Analysis and Adaptation (SODAA) tool[6]. This tool analyzes the Surface Management System[7] (SMS) generated log files, which contain data from multiple sources, including air carriers, the Enhanced Traffic Management System (ETMS), and Airport Surface Detection Equipment, Model X (ASDE-X). SODAA allows the user to extract large amounts of information about surface traffic at DFW airport. The type of data extracted is explained below. SMS log files are stored in twenty-four hour datasets, of which fifteen were selected from the year 2008 for analysis. The days selected were: April 6, April 8, April 14, April 15, April 16, April 19, April 26, April 29, May 1, May 24, May 26, June 11, July 8, July 10, and July 11. All days experienced clear weather, with the majority of the airport configuration in south flow, where arrivals used runways 13R, 18R, 17C, and 17L, and departures used runways 18L, 17R, and 13L. Clear weather was assumed to be visibility greater than 10 statute miles, dry runways, and cloud cover representing visual meteorological conditions (VMC).

Data collected was either processed directly within SODAA, or post-processed using Matlab. A thorough description of the fields extracted from SODAA, as well as the values derived from these fields is given in the next sub section. A select number of SODAA fields used for this analysis are grouped into four different categories: flight identification, position and time, route, and efficiency. Flight identification includes call sign, category (arrival or departure), aircraft type, weight class, and engine type. Position and time data includes queue time, runway occupancy time, OFF time[‡], ON time[§], SPOT time[**], x and y coordinates, and speed output at one second intervals between the spot and runway. Route data includes the runway, taxi route, spot, fix, and destination airport.

B. Calculating Taxi Time and Influencing Factors

Taxi time was derived by subtracting the SPOT time from the OFF (take-off) time for departures, or the ON (landing) time from the SPOT time for arrivals. Queue time was derived using the aircraft's position data, and the runway assignment. A queue entry point was extracted from X and Y coordinates, depending on the runway involved. When the flight passes this point, it was considered to be in the queue. Queue time was then calculated by subtracting queue entry time from OFF time. The portion of the taxi until the queue entry for departures is called the adjusted taxi time. It was calculated as the SPOT time subtracted from the queue entry time. Arrival flights do not have a queue time or adjusted taxi time.

The service rate was defined as the number of arrivals and departures that were activated (enter the surface) 10, 20, or 30 minutes before the flight in question is activated. SPOT times were used as activation times for departures, while ON times were used as activation times for arrivals. For example, the departure service rates for AAL123 is the number of departure flights that were released from the spot 10, 20, or 30 minutes prior to AAL123 being released. The service rate was also calculated for the number of arrivals and departures that were de-activated (exit the surface of the airport), but their minimal influence on taxi time resulted in their disregard.

Aircraft position and speed data was used along with runway assignment to determine the number of stops and where the stops occurred for each flight. A stop was recorded when a flight's speed dropped below the threshold speed of three knots for at least five seconds, then increased above three knots. This threshold speed was determined by observing position and speed data from SODAA, in which an aircraft could display a speed up to three knots without any significant change in position, which meant it was not moving. This can be seen in Figure 2, where the position coordinates are shown for an aircraft that is conducting a stop. The number of stops was then split into queue stops and taxiway stops depending on the position data when the stop occurred. Arrival flights only incurred taxiway stops. The stopped duration for each flight was calculated as the time the aircraft's speed was below three knots. Position data was used to classify time stopped as taxiway time stopped and queue time stopped for departure flights. Arrival flights only had time stopped on the taxiway. As an example, Figure 3 is the speed profile of a flight that stopped twice for a total of 67 seconds.

Surface traffic count was found using position data and flight ID's for all aircraft. Using position and time data, the number of flights that were active on the taxiway or in the queue was counted. This value was calculated every sixty seconds along the flight's route from the time it left the spot to when it departed, or from the time it touched down to the time it reached the spot. The average of these counts was used as the surface count for that flight. Position and speed data allowed the breakdown of surface count into regions (See Figure 1). Total surface count was the sum of aircraft counts in each region. The region that each flight passed through was also recorded to keep track of which parts of the surface affected the flight.

Using each flight's taxi route, runway and spot, the distance taxied for each flight was estimated. Using actual position data provided very noisy results for taxi distance, so taxi route information was coupled with SODAA intersection coordinates to calculate taxi distance. The taxi route was broken down into intersections, and the corresponding coordinates were found. Euclidean distance was then calculated between the coordinates to provide taxi distance traveled. The runway and spot assignments provided a beginning and end intersection to complete the

[‡] Time of beginning of movement on runway which continues along runway to flight speed

[§] Time at which aircraft crosses a runway threshold and slows to taxi speed

[**] Time at which the flight crosses a spot

taxi distance analysis. Although “smoothing the data” was another option for calculating distance, incomplete position data for some flights would still result in erroneous distance calculations; hence the use of the above technique. The number of turns for each flight was also calculated using the taxi route, which included intersections where aircraft leave one taxiway for another resulting in a turn. The number of turns was calculated by parsing the taxi route, adding spot and runway assignment, and determining how many turns was necessary to maneuver on the number of taxiways specified, while moving from spot to specified runway.

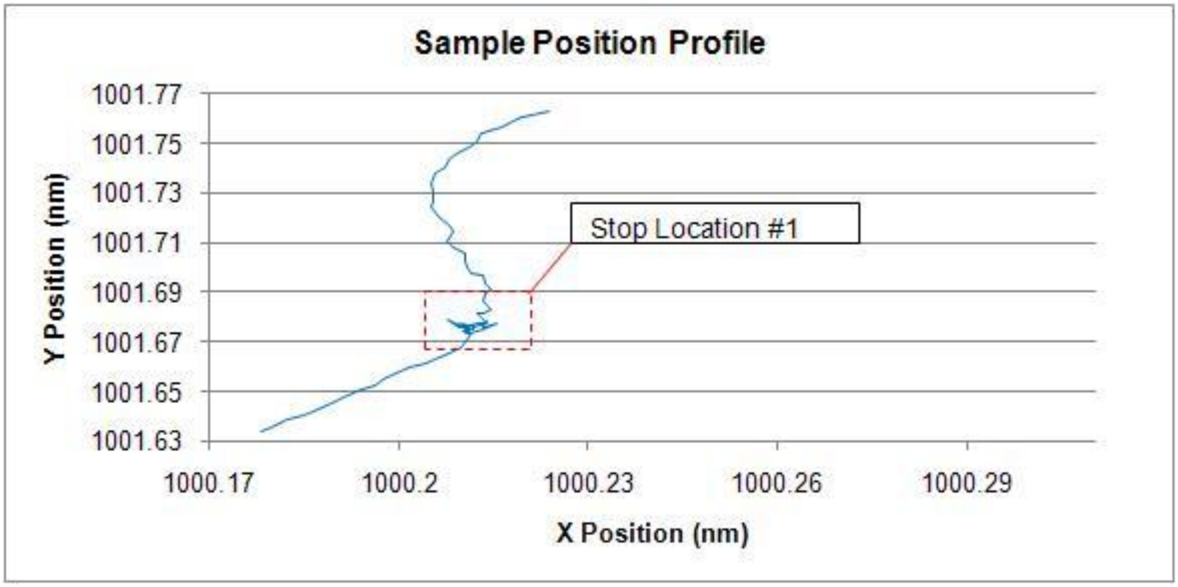

Figure 2. Sample Position Profile

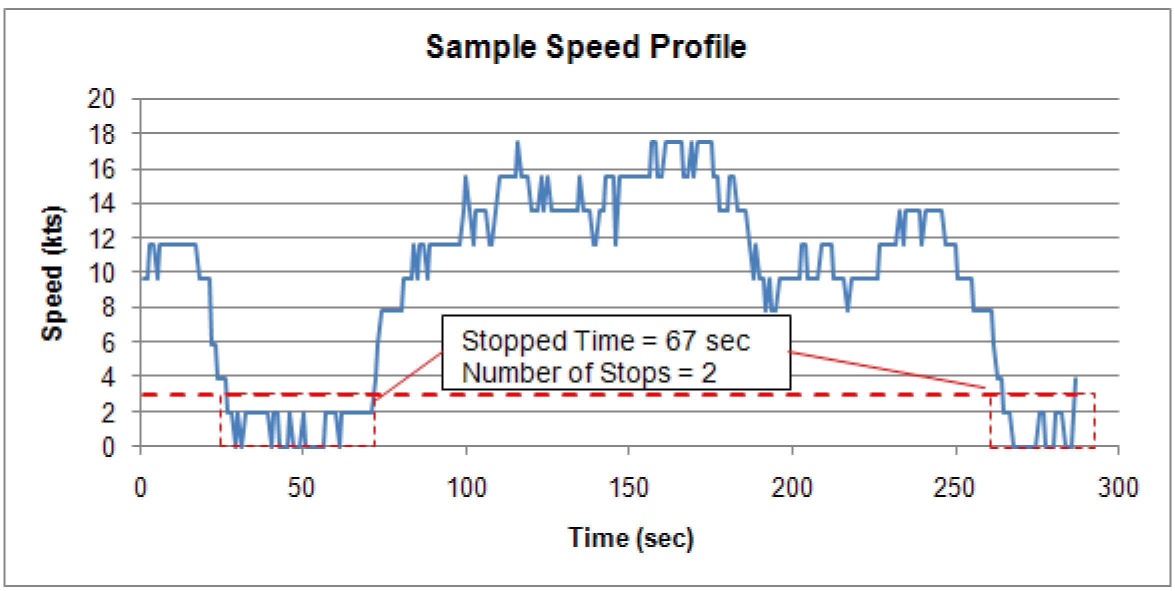

Figure 3. Sample Speed Profile for Two Stops

C. Estimation Models

This section details the models for estimating taxi time as a function of various aggregate parameters. The analysis is based on two assumptions:

- The weight class of the aircraft has not been treated as a relevant parameter. Almost 85% of the operations at DFW are of weight class Large[††], and the dataset collected contained insufficient amounts of other weight classes to consider their effects. Further, given the numerous amount of Large aircraft, it is possible that weight class may not be a significant determinant of taxi time.
- For departure aircraft, the taxi time does not include the time spent in the departure queue. Departure queue operations include decision making based on a variety of factors including ground delay restrictions, weather related miles-in-trail restriction and others. The current dataset does not capture these decision making factors. It is plausible that such factors influence the time from spot to queue as well, but the results in following paragraphs show that such factors do not influence spot to queue taxi times considerably.

Two functional forms were used to evaluate the relationship as shown below in equations (1) and (2), where equation (1) represents a linear form and equation (2) represents a log-linear form. The two functional forms represent the cases when the effect of the parameters is additive and multiplicative respectively. Also, since both the functional forms can be represented linearly as shown below, ordinary least squares is used for estimation. The simplistic functional forms were chosen due to a lack of past studies in this regard; presence of such a relationship would be a motivating factor for testing advanced forms. There is no constant included in either functional form because a zero distance traveled should result in a zero taxi time. An incremental approach was used for each model, where initially estimation was done including all identified independent variables, and then the non-significant variables were discarded and the model was re-estimated.

For the sake of clarity, a naming convention was defined for each model as follows: the first letter in the name defines the functional form (L for Linear, M for Multiplier or log-linear). The second letter denotes whether the total surface count variable (TotalTraffic) was used, represented by T or the regional variables (RegionTraffic and NonRegionTraffic) were used, represented by R. If stop variables were used, these letters are followed by –S. In the end, a * denotes if only significant variables were used. Thus, model LT-S* is the linear form including all traffic and stops, with significant variables only, whereas model MR denotes the log-linear form including regional traffic only with no stop variables and non-significant variables are present too. The results from the linear functional form are presented in the results along with those from the log-linear functional form.

$$\begin{aligned}\text{taxi time} = &\sum \text{coefficient} \times \text{non-zero independent variable} \\ &+ \sum \text{coefficient} \times \text{positive independent variable} \\ &+ \text{coefficient} \times \text{dummy variable}\end{aligned} \tag{1}$$

$$\begin{aligned}\ln(\text{taxi time}) = &\sum \text{coefficient} \times \ln(\text{non-zero independent variable}) \\ &+ \sum \text{coefficient} \times \text{positive independent variable} \\ &+ \text{coefficient} \times \text{dummy variable}\end{aligned} \tag{2}$$

III. Results

A. Comparative Statistics of some Taxi Characteristics

This section details comparative statistics of certain taxi trajectory characteristics. Besides providing the reader with an overview of the relevant quantities used in analysis, this section also compares arrival and departure operations on the surface and illustrates the difference between the two. Figure 4 shows a histogram of average taxi speed for arrival and departure aircraft (Figure 4(a) shows arrivals and Figure 4(b) shows departures) found using SODAA in which over 17,000 flights are shown. The average speed is the total distance traveled by the aircraft on the surface divided by the time taken; this *includes* the queue waiting time for departure aircraft and the runway

[††] Large aircraft are aircraft with a take-off gross weight between 41,000 and 255,000 pounds

crossing waiting time for arrival aircraft. The figure shows that for routes with similar distances, departure aircraft have longer taxi times than arrival aircraft.

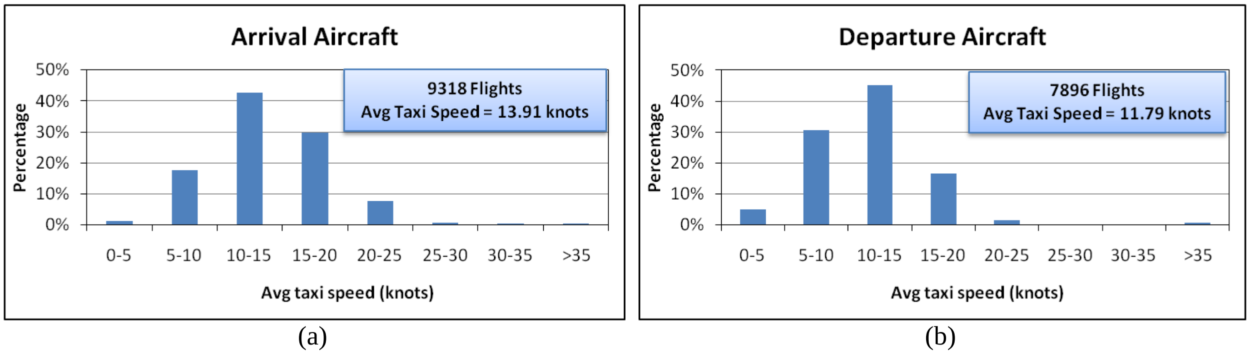

Figure 4. Average Taxi Speed for Arrival and Departure Operations

Figure 5 is a histogram of the amount of time flights are stationary on the taxiway as defined before. The difference between arrival aircraft and departure aircraft is very stark here: almost 60% of the arrivals stop for less than ten seconds as compared to almost 75% of the departure aircraft stopping for more than ten seconds.

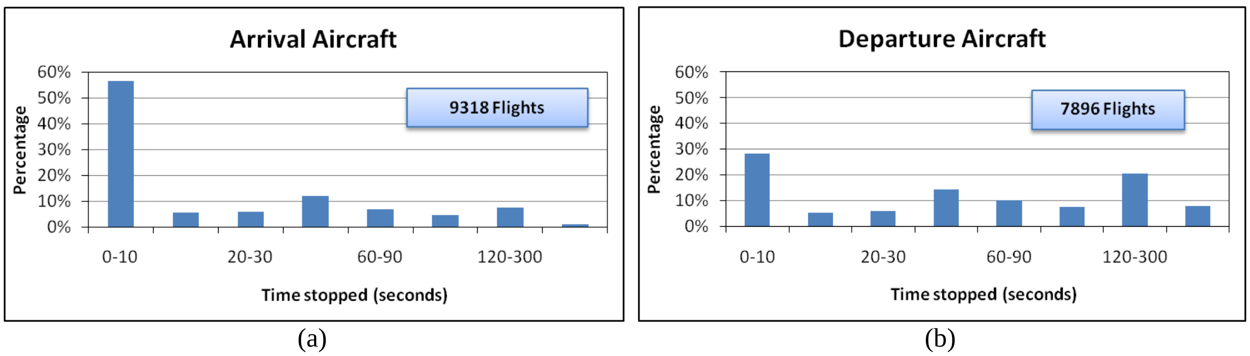

Figure 5. Time Stopped for Arrival and Departure Aircraft

A potential reason for the large amount of time stopped for departure aircraft was holding in the departure queue at the runway and waiting for clearance. To compare stopping time at the runway queues and on taxiways, Figure 6 includes histograms for departure aircraft stopping time for queues and taxiways. Almost 80% of the departure aircraft stopped for less than ten seconds on the taxiways, whereas the time stopped in the queues was higher and varied. One potential reason could be that some take-off time and sequence related decisions due to weather at destination airports (EDCT, miles-in-trail and others) were made at the departure queues rather than before pushback or on the taxiway.

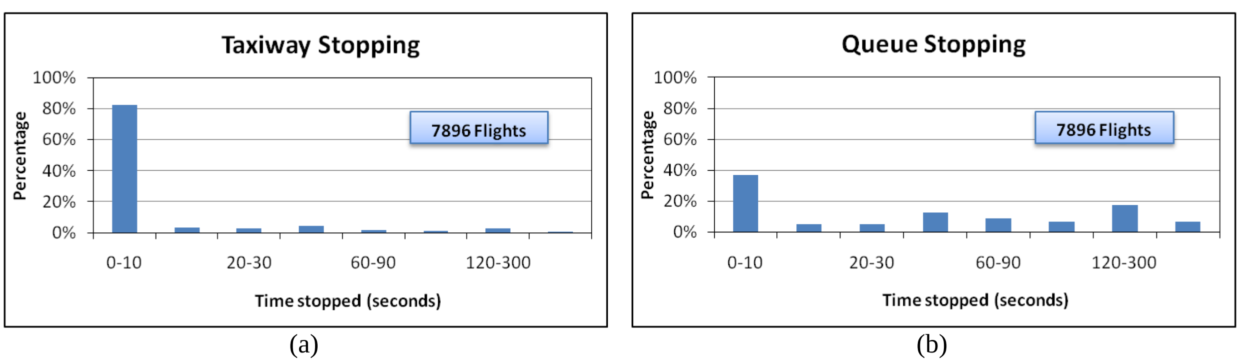

Figure 6. Departure Aircraft Time Stopped on Taxiways and Departure Queues

B. Estimating Taxi Time Relationships

Table 1 below lists all the variables extracted from SODAA including the factors explored in estimation, their names and meaning. The variables are separated based on their usage in models as described in equations (1) and (2). All the time related variables are in seconds.

Table 1. Variables Used in Model Estimation, Separated Based on Usage in Models

Name	Meaning
Dependent Variable	
TaxiTime	time on taxiways (seconds)
Independent Variables (non-zero)	
Turns	number of turns in the taxi trajectory
Distance	distance traveled by aircraft while taxiing (in x 10^{-3} nautical miles)
Independent Variables (greater than or equal to zero)	
Dep 0-10	number of departure aircraft released from spot in the 10 minutes before aircraft activation
Dep 10-20	number of departure aircraft released from spot between 10 to 20 minutes before aircraft activation
Dep 20-30	number of departure aircraft released from spot between 20 to 30 minutes before aircraft activation
Arr 0-10	number of arrival aircraft landings in the 10 minute before aircraft activation
Arr 10-20	number of arrival aircraft landings between 10 to 20 minutes before aircraft activation
Arr 20-30	number of arrival aircraft landings between 20 to 30 minutes before aircraft activation
TaxiStops	number of stops on the taxiway
TaxiStopTime	stopping time on taxiways (seconds)
RegionTraffic	average surface traffic in pertinent regions while aircraft taxies
NonRegionTraffic	average surface traffic outside pertinent regions while aircraft taxies
TotalTraffic	average total surface traffic while aircraft taxies = RegionTraffic + NonRegionTraffic
Independent Variable (Dummy)	
ArrState	dummy variable, 1 if aircraft is an arrival, 0 if aircraft is a departure

The service rate variables (Dep 0-10, Dep 10-20, Dep 20-30, Arr 0-10, Arr 10-20 and Arr 20-30) include the number of arrival and departure aircraft *entering the taxiways* in the last thirty minutes (parsed out in ten minute intervals) before the flight in question is activated. These service rates were chosen because their values can be found fairly easily by the controller and would have already occurred when the flight is ready to be activated. Model estimation was done using runs with and without the number of stops and corresponding time stopped as variables. These variables will not be known ahead of time by the controller, but may be estimated to some degree. Also, the inclusion of these variables in estimation provides insight into the contributive effect of each additional stop and additional stop time, besides producing better estimation fit.

1. *Linear Models*

The variables and coefficients of the first linear model, denoted as Model LR, are shown in Table 2. Non-significant values are darkened.

Table 2. Estimation Results for Model LR

Variable	Coefficient	Std. Error
Turns	3.258	0.781
Distance	0.152	0.002
ArrState	110.452	1.830
Dep 0-10	0.142	0.358
Dep 10-20	-1.006	0.352
Dep 20-30	-1.742	0.343
Arr 0-10	1.779	0.351
Arr 10-20	-0.251	0.357
Arr 20-30	-0.345	0.340
RegionTraffic	11.815	0.450
NonRegionTraffic	-1.521	0.507

Note: Coefficient units are seconds/variable. For example, (2 turns)(3.258 sec/turn) = 6.516 sec*

As expected, the arrival dummy's coefficient is significant, and the coefficient value states that the arrival aircraft typically spent 110 more seconds taxiing than a comparable departure aircraft, when the time spent in the departure queue for the departure aircraft was excluded. The results show an adjusted coefficient of determination (R^2) value of 0.91, and for a better fit the use of number of stops and total stop time as independent variables in the estimation was explored. Results from the estimation of this model (labeled LR-S) are shown below in Table 3. Again, the non-significant estimates are darkened. There is a noticeable increase in the adjusted R^2 values, and the taxi stop time as well as number of stops is estimated as significant. Although the inclusion of these variables in the estimation process raises doubts about the real time usage of such a model (since the number of stops and stop time is not known a priori), significant estimates imply that their effect cannot be neglected in model building. However, for a future system that seeks to minimize stops, inputs pertaining to these variables would be minimal, and thus would not result in high errors.

Model LR-S* builds on the above by removing the non-significant variables and re-estimating the model. The results are presented in Table 4, and all estimates are significant here with a high adjusted R^2 value. The model provides significant insight into surface operations, showing that each turn added about three seconds to the taxi time, while traversing a nautical mile took about 157 seconds. Every aircraft in the region that the flight will taxi through added about six seconds to the taxi time. Taxi stops increased the taxi time by about nine seconds per stop. Each arrival aircraft spent almost 83 more seconds in similar conditions than a departure aircraft when departure queue time was excluded.

The negative coefficients for some of the service rate variables counter expectation, since increased inflow in the system should result in more congestion and higher taxi times. However, the standard errors for these estimates are high, which questions the estimates themselves. One potential reason for this could be the correlation in inflow and traffic, since more inflow would mean more traffic. Also, the coefficient for traffic outside regions of taxi route is significant yet negative, which resists physical interpretation since additional aircraft in a non-pertinent region should not affect time. To resolve this, models with total surface traffic are also considered. Estimation results from model LT-S are given in Table 5; this includes all variables, with the non-significant estimates darkened. In model LT-S*, the non-significant variables were removed and the model was re-estimated, with the results in Table 6. The estimated coefficients in LT-S* are not very different from model LR-S*, with each turn resulting in two seconds more taxi time, and each nautical mile adding about 168 seconds. The coefficient for total traffic states that each additional aircraft on the surface added about two seconds to the taxi time. However, the standard error in the estimate is much higher as compared to the regional traffic estimate in model LR-S*. The resulting adjusted R^2 value, however, is comparable to model LR-S*.

Table 3. Estimation Results for Model LR-S

Variable	Coefficient	Std. Error
Turns	2.960	0.462
Distance	0.157	0.001
ArrState	82.641	1.097
Dep 0-10	-0.954	0.212
Dep 10-20	-0.748	0.208
Dep 20-30	-0.916	0.203
Arr 0-10	0.878	0.208
Arr 10-20	0.136	0.211
Arr 20-30	-0.078	0.201
RegionTraffic	6.491	0.268
NonRegionTraffic	-3.554	0.300
TaxiStopTime	1.068	0.008
TaxiStops	9.300	0.491
Adjusted R^2	**0.968**	

Table 4. Estimation Results for Model LR-S*

Variable	Coefficient	Std. Error
Turns	2.975	0.457
Distance	0.157	0.001
ArrState	82.668	1.090
Dep 0-10	-0.945	0.211
Dep 10-20	-0.739	0.205
Dep 20-30	-0.918	0.201
Arr 0-10	0.909	0.181
RegionTraffic	6.493	0.268
NonRegionTraffic	-3.542	0.298
TaxiStopTime	1.067	0.008
TaxiStops	9.301	0.491
Adjusted R^2	**0.968**	

Note: Coefficient units are seconds/variable. For example, (2 turns)(2.96 sec/turn) = 5.92 sec*

Table 5. Estimation Results for Model LT-S

Variable	Coefficient	Std. Error
Turns	2.502	0.469
Distance	0.168	0.001
ArrState	85.571	1.107
TaxiStopTime	1.071	0.008
TaxiStops	9.600	0.498
TotalTraffic	1.941	0.182
Dep 0-10	-1.047	0.215
Dep 10-20	-1.044	0.211
Dep 20-30	-1.203	0.206
Arr 0-10	0.112	0.208
Arr 10-20	-0.115	0.214
Arr 20-30	-0.265	0.204
Adjusted R^2	**0.967**	

Table 6. Estimation Results for Model LT-S*

Variable	Coefficient	Std. Error
Turns	2.367	0.455
Distance	0.168	0.001
ArrState	85.343	1.094
TaxiStopTime	1.071	0.008
TaxiStops	9.596	0.498
TotalTraffic	1.910	0.180
Dep 0-10	-1.095	0.211
Dep 10-20	-1.119	0.205
Dep 20-30	-1.269	0.199
Adjusted R^2	**0.967**	

Note: Coefficient units are seconds/variable. For example, (2 turns)(2.502 sec/turn) = 5.004 sec*

2. Log-linear Models

This section includes the results from estimating the log-linear models. The relevant variables which were used in the logarithmic form during estimation are denoted by "ln(variable)". The variables and coefficients of the first log-linear model with regional and non-regional traffic separated out (labeled as Model MR) are shown below in Table 7. As before, non-significant values are darkened, and system inflow beyond 20 minutes before activation is not significant. The amount of traffic outside relevant region is also not significant. Model MR* is a result of removing the non-significant variables from model MR and re-estimation, with the results in Table 8. Although the coefficient for number of turns is negative, the value is very small compared to one, which means that there was a very small increase in taxi time with each additional turn. Arrival aircraft were predicted to taxi about 1.6 ($e^{.47}{\sim}1.6$) times longer than departure aircraft. The region count did not affect taxi time very significantly for this log-linear model. Service rates once again seemed to have a small effect on taxi time. The use of a log-linear model, even without stops, improves the adjusted R^2 value to 0.997. The standard error in the estimates of service or inflow rates as well as number of turns is relatively high compared to the estimate itself.

Table 7. Estimation Results for Model MR

Variable	Coefficient	Std. Error
ArrState	0.471	0.005
Dep 0-10	0.004	0.001
Dep 10-20	0.002	0.001
Dep 20-30	-0.001	0.001
Arr 0-10	0.006	0.001
Arr 10-20	0.002	0.001
Arr 20-30	0.002	0.001
ln(Turns)	-0.015	0.008
ln(Distance)	0.723	0.002
RegionTraffic	0.033	0.001
NonRegionTraffic	0.000	0.001
Adjusted R^2	**0.997**	

Table 8. Estimation Results for Model MR*

Variable	Coefficient	Std. Error
ArrState	0.472	0.005
Dep 0-10	0.004	0.001
Dep 10-20	0.002	0.001
Arr 0-10	0.007	0.001
Arr 10-20	0.003	0.001
ln(Turns)	-0.014	0.007
ln(Distance)	0.723	0.002
RegionTraffic	0.033	0.001
Adjusted R^2	**0.997**	

Note: Coefficient values are found by (ArrState = $e^{0.471}$ = 1.6), which means Arrivals taxi 1.6 times longer, in seconds, than departures

To gauge the effect of stops, the log-linear functional form was tested with the inclusion of stop time and taxi stops. The variables and coefficients for this model (labeled MR-S) are shown below in Table 9. Again, non-significant values are darkened, and some of the inflow variables are not-significant. In model MR-S* (Table 10) these are removed and the model is re-estimated which shows that taxi distance was the dominating factor, more than doubling the taxi time for each nautical mile traveled. There is a small increase in the adjusted R^2 value, and the coefficients for stops and time stopped are significant. The variable for traffic in non-relevant regions is also significant, but the very small negative coefficient implies that each additional aircraft in non-relevant region results in a very small increase in taxi time.

Table 9. Estimation Results for Model MR-S

Variable	Coefficient	Std. Error
ArrState	0.410	0.005
ln(Turns)	-0.024	0.006
ln(Distance)	0.730	0.002
RegionTraffic	0.022	0.001
TaxiStops	0.049	0.002
TaxiStopTime	0.002	0.000
Dep 0-10	0.001	0.001
Dep 10-20	0.002	0.001
Dep 20-30	0.000	0.001
Arr 0-10	0.004	0.001
Arr 10-20	0.003	0.001
Arr 20-30	0.002	0.001
NonRegionTraffic	-0.006	0.001
Adjusted R^2	**0.998**	

Table 10. Estimation Results for Model MR-S*

Variable	Coefficient	Std. Error
ArrState	0.410	0.005
ln(Turns)	-0.024	0.006
ln(Distance)	0.730	0.002
RegionTraffic	0.023	0.001
TaxiStops	0.049	0.002
TaxiStopTime	0.002	0.000
Dep 10-20	0.002	0.001
Arr 0-10	0.004	0.001
Arr 10-20	0.003	0.001
Arr 20-30	0.002	0.001
NonRegionTraffic	-0.006	0.001
Adjusted R^2	**0.998**	

Note: Coefficient values are found by (ArrState = $e^{0.410}$ = 1.51), which means Arrivals taxi 1.51 times longer, in seconds, than departures

A summary of the results from the different models and their setup is shown in Table 11.

Table 11. Model Summary Including Variables Used

Model #	Type	Surface Count	Stops Included	Variables	Adjusted R^2
LR	Linear	Regional	No	All	0.910
LR-S	Linear	Regional	Yes	All	0.968
LR-S*	Linear	Regional	Yes	Significant	0.968
LT-S	Linear	Total	Yes	All	0.967
LT-S*	Linear	Total	Yes	Significant	0.967
MR	Log-linear	Regional	No	All	0.997
MR*	Log-linear	Regional	No	Significant	0.997
MR-S	Log-linear	Regional	Yes	All	0.998
MR-S*	Log-linear	Regional	Yes	Significant	0.998

C. Testing Estimated Models on an Independent Dataset

Only five models with significant estimates were selected to estimate the taxi times for an independent twenty four hour dataset from May 24, 2008. This dataset was excluded from the model-estimation in the earlier section so it could be used for testing the models. The weather conditions and configuration are similar to the datasets collected to find the relationships between taxi time and aggregate factors of airport traffic movement. Selections included Model LR even though it includes non-significant variables, to demonstrate the importance of including stop situations in the linear models. The absolute difference in taxi times was used to compare the models. Figure 7 shows a histogram of the absolute difference in taxi times between the three linear models (LR, LR-S* and LT-S*), while Figure 8 plots the difference between the two log-linear models (MR* and MR-S*). It can be seen that the linear models LR-S* and LT-S* perform relatively better than LR.

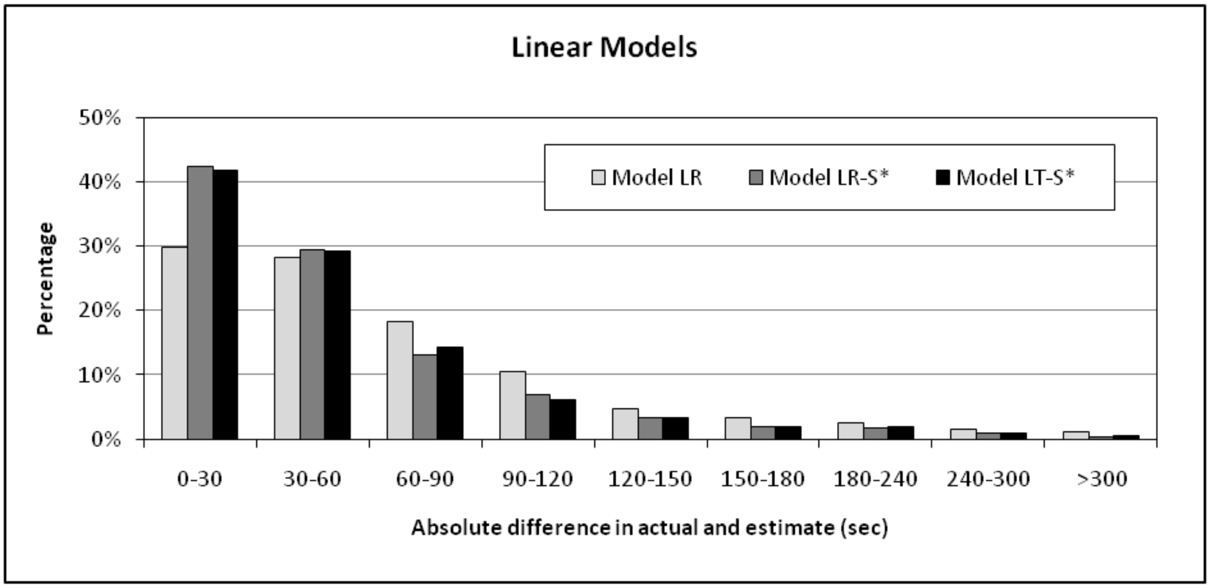

Figure 7. Absolute Difference in Taxi Times for Linear Models

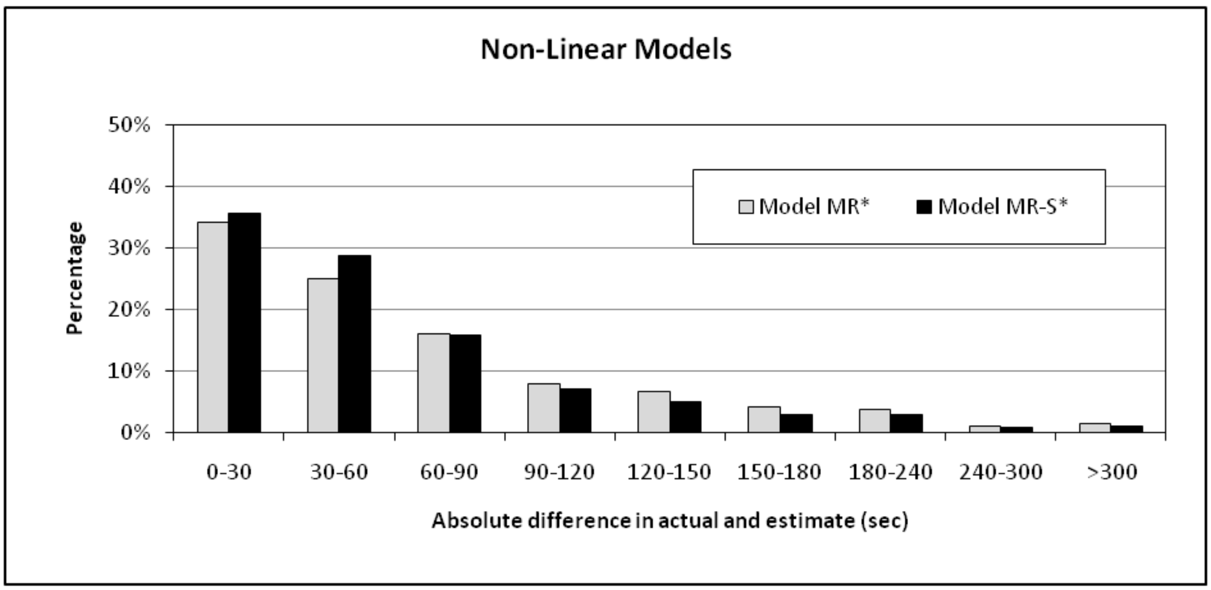

Figure 8. Absolute Difference in Taxi Times for Log-Linear Models

A comparison of the percentage of taxi times estimated within one minute of the actual for these models can be seen below in Table 12. The one minute limit was selected since the separation requirements for successive operations on a runway result in a minimum one minute gap. Thus, if "predictions" from the above models are used to facilitate some metering mechanism for taxiway usage, a minimum of one minute accuracy would be required. The linear models LR-S* and LT-S* perform better, with each estimating greater than 70% of the taxi times within one minute of the observed data. The log-linear models did not perform as well, with Model MR-S* performing the best at just under 65% of taxi times within one minute of observed data.

The model LR includes all variables, significant or not. As seen in Table 12, the inclusion of stop situations in linear models results in almost 12% increase in "better" estimation, whereas the effect is much more subdued in log-linear models, with an increase of only about 5%. This underscores one of the key reasons for unsatisfactory estimation: the choice of an appropriate functional form. Exploring other functional forms is a direction for future work beyond this paper, and would involve more involved processes like maximum likelihood estimation, besides exploring a wide variety of possible functions. Another reason for poor estimation is the correlation between independent variables. The number of aircraft entering the system in the last 10 minutes would be correlated with the amount of traffic during taxiing, and the functional forms used in here do not account for such correlation. Lastly, the dataset selected for testing itself might have some peculiarities absent from the dataset used from estimation.

Table 12. Percentage of Taxi Times within 1 minute

Model	Percentage
Model LR	58.11%
Model LR-S*	71.86%
Model LT-S*	71.07%
Model MR*	59.14%
Model MR-S*	64.55%

IV. Conclusion

A relationship between taxi time and various aggregate traffic factors was tested using two functional forms: linear and log-linear. The results indicate that distance traveled on the surface, number of stops, and surface traffic (region and total) have a significant effect on taxi time. Both linear and log-linear functional forms fit the data with R^2 values greater than 0.95. Although not designed for prediction, the models were nevertheless tested over an independent dataset recorded on May 24, 2008. Linear models performed better than log-linear models, with more than 71% of the taxi times estimated within one minute of the observed data. Log-linear models performed slightly lower, with just under 65% of the taxi times predicted within one minute of the observed data. Correlation between the aggregate factors could be a reason for poor estimation. Also, only simple functional forms are used in this paper. More complex functional forms requiring maximum likelihood estimation need to be tested to model surface factors and their effect on taxi times. Future work would include testing such functional forms for a better estimation as well as addressing correlation between influencing factors.

The observed relationship between aggregate traffic and taxi time can be used as a mechanism for metering taxiway access, leading to better surface operations especially under heavier traffic demand. Even in the absence of complex mechanism for such metering, it is possible to visualize a scenario where the surface controller uses the above relationship as simple rules to anticipate changes in surface efficiency. Of course, such relationships are very airport specific and would require substantial analysis of current operations at the relevant airport. Even though the relationships could be different, a complete lack of such a relationship would be surprising.

Acknowledgements

The authors would like to thank the following staff at NASA Ames Research Center for their valuable comments to the paper: Leonard Tobias, Waqar Malik, and William Chan.

References

[1] Joint Planning and Development Office (JPDO), "Concept of Operations of the Next Generation Air Transportation System," Version 2.0, June 13, 2007

[2] Brinton, C., Wood, B., Engelland, S., "Microscopic Analysis of Airport Surface Sequencing," AIAA 26th Congress of International Council of the Aeronautical Sciences (ICAS) including 8th AIAA Aviation Technology, Integration, and Operations, Anchorage, Alaska, Sep. 14-19, 2008

[3] Atkins, S., Brinton, C. and Jung, Y. "Implication of Variability in Airport Surface Operations on 4-D Trajectory Planning," AIAA 26th Congress of International Council of the Aeronautical Sciences (ICAS) including 8th AIAA Aviation Technology, Integration, and Operations, Anchorage, Alaska, Sep. 14-19, 2008

[4] Howell, D., Flanders, I., Shema, S., "Using Surface Demand Trends to Evaluate Multiple Airport Surface Initiatives," AIAA 7th Annual Aviation Technology, Integration, and Operations Conference (ATIO), Belfast, Northern Ireland, Sept. 18-20, 2007

[5] Pledgie, S., Atkins, S., Brinton, C., "Methods for Aircraft State Estimation from Airport Surface Surveillance," Mosaic ATM, AIAA Guidance, Navigation, and Control Conference (GNC), Chicago, Illinois, Aug. 10-13, 2009

[6] SODAA, Surface Operations Data Analysis and Adaptation tool, Software Package, Ver. 1.8, Mosaic ATM, Leesburg, VA, 2008.

[7] Raytheon ATMSDI Team, "SMS Final Report," May 2004