Analyzing and Decomposing Taxi Times and Predicting Taxi out Times

Yu Zhang and Arjun Chauhan
Department of Civil and Environmental Engineering
University of South Florida
Tampa, Florida, USA
yuzhang@eng.usf.edu

Xing Chen
Federal Aviation Administration
Washington DC, USA
Xing.chen@faa.gov

***Abstract* — This paper proposes a set of regression equations to model the taxi-out and taxi-in times at airports. The estimated results can be used to calculate the nominal taxi times, which are essential measures for evaluating the taxiing delays at airports. Given the outcomes of the regression model, an iterative algorithm is developed to predict taxi times with inputs such as gate out times, landing times, and runway capacities. Case study at LGA shows that the proposed algorithm demonstrates a higher accuracy in comparison to other algorithms in existing literature.**

Keywords-taxi time delay; nominal taxi times; predicting taxi-out time; iterative algorithm.

I. INTRODUCTION

The rise of urbanization has taken its toll on the airline industry among many others. There has been consistent increase in airline traffic from the time it began. Today there are about 7000 flights in America's skies during the peak hours. This is despite a slump in air traffic recently due to the global market meltdown. The air traffic has still been up when compared to the periods before the recession. This rise in airline traffic has seen major delays in the National Airspace System (NAS). A large percentage of flight delay is due to ground holding [(1)] and ground transit which includes taxiing delay. Taxi times are the times spent by an aircraft between rolling from a gate to the end of a runway where it takes off or from the entrance of taxiways to a gate after it lands on a runway. Taxiing-in and out are major parts of arrival and departure processes. Considering the distribution of delays experienced by a flight, taxi out delay contributes to 26 percent of the total. According to BTS, 2007 has been a year of the highest taxi times recorded that surpassed the previous peak in the year 2000[(2)]. The average block times between busy city pairs in the U.S. increased accordingly, for example, according to Air Transport Association (ATA), in New York LaGuardia (LGA) – Ronald Reagan Washington National (DCA) route segment, the average block time grew by nine minutes from 1995 to 2005[(11)]. Longer taxi times have elevated the direct operating and maintenance costs as well as negative environmental impacts in terms of amplified noise and augmented air pollution on and around the airport.

To mitigate delay problems, the FAA implemented the Collaborative Decision Making (CDM) approach in 1998. The CDM in the US is intended towards improving air traffic flow issues in the National Airspace System (NAS) through exchange of information among the air traffic flow managers, air traffic controllers, and airlines. In the US, the initial focus of the CDM was the Ground Delay Program Enhancements (GDP-E) where the airlines share flight cancellation and reordering information with the Air Traffic Control System Command Center. The users of the NAS also use CDM tools to share information on safety and efficiency among themselves. The CDM concept applied to some EU airports is known as Airport CDM[(11)] (A-CDM). The focal point of A-CDM is to bring together the major airport partners like air traffic controllers, aircraft operator, ground handlers and share data in a clear manner. This becomes significant to achieve a common situational understanding consequently leading to better decision making processes.

Presently, the Next Generation Air Transportation System (NextGen) is under way, the objective is to improve the NAS to meet future demand, avoid congestion, and make the skies safer. NextGen[(6)] suggests using various technologies, equipment, and procedures to enhance pilots' control over flight paths while the controllers on the ground focus more on traffic flow management. NextGen looks to implement new tools that are being developed to help manage aircraft flow at airports in order to mitigate taxiing delays, reduce engine run times and consequent environmental impact. Such new tools require a better understanding of the taxi times, taxiing delays, and also call for a way to accurately predict taxi times. Accurate prediction of departure taxi times are essential and help airlines manage push back times, obtain and pass on delay information to destination airports. Correct prediction is a key component of the CDM operations and leads to better gate management and reduced arrival and departure delays. The Air Traffic Control (ATC) will benefit as well via improved demand forecasts for airports and en-route air sectors.

This study contains two parts. In the first part, a set of linear models are established to model the taxi-in and taxi-out times. Besides offering inputs for the predicting model in the second part, the set of linear models can be used to calculate the nominal taxi times, which are essential measures that can be used to evaluate the taxiing delays at the airports. In the second part, an iterative algorithm is proposed to predict the taxi-in and taxi-out time with the outcomes from the regression models and other inputs. In comparison to other existing taxi time predicting model, the outcomes of the case study with our model provide higher accuracy and reliability.

II. LITERATURE REVIEW

The existing model for estimating unimpeded taxi times recorded in the Aviation System Performance Matrix (ASPM) database is developed by Kondo[5] based on two linear equations, one for taxi-in, the other for taxi-out, while containing both taxi-in and taxi-out queue lengths. Given the actual flight information, such as, actual departure and wheel-off times, Kondo sets up bins for each minute of a single day and count how many departing aircraft ahead of one flight at the queue entry time (gate out time). The number of aircraft ahead is considered as the departure queue length for that flight. Arrival queue length can be obtained in a similar way but considering wheel-on and gate in times. For each group, defined according to carrier and season, the taxi-out time is then modeled as the linear combination of an intercept, weighted taxi-out queue length, and weighted taxi-in queue length, as well as the taxi-in time with a different set of coefficients. Given the recorded data, the intercept and weights (coefficients) can be regressed with Ordinary Least Square method. Assuming the interested flight is the only aircraft moving in the taxiway systems, the nominal taxi-out times are calculated with the regression results and by setting the departure queue length as 1 and arriving arrival queue length as 0. Similarly, the nominal taxi-in time is obtained by setting the number of arriving queue length to be 1 and departing queue length to be 0 in the equation. This model captured the major factor contributing to taxi times, the queue lengths of arrival and departure flights. However, it did not consider other factors such as runway configurations, weather impact, and others.

Causal factors identified in Idris et al's paper include runway configuration, airline/terminal location, departure demand, departure queue size, weather, and downstream restrictions. They stated that the runway configuration determines the flow of aircraft at the airport, presents the level of interaction between the flows, and restricts the capacity of arrivals and departures. Idris et al also discussed weather and downstream restrictions in view of the fact that adverse weather greatly reduces the capacity of the airport. They suggested another way of calculating the arrival and departure queue length, accounting for the passing of aircraft, which is shown in Fig. 1.

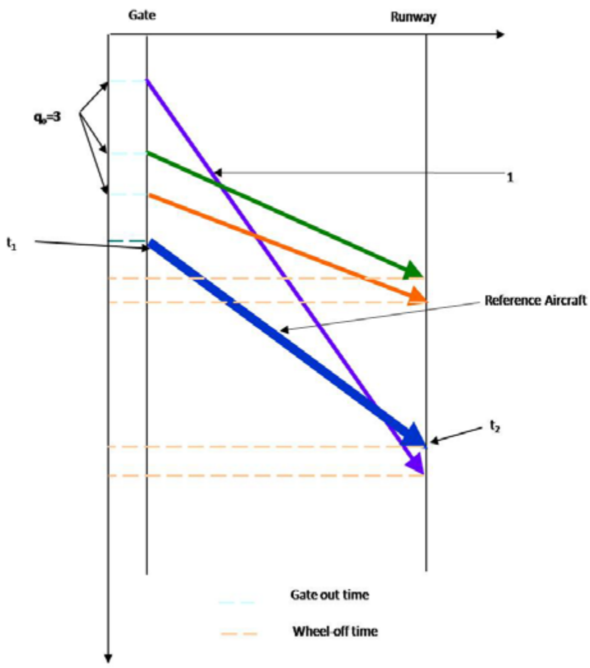

FIGURE. 1 Queue Length Calculation

Fig. 1 shows four aircraft taxing-out from the gate and taking off. The reference aircraft leaves the gate at a time t_1 and takes-off at a time t_2. The taxi out duration of the reference aircraft is t_2 - t_1. There are three aircraft that have a gate out time before t_1. However, aircraft 1 takes off at a time after t_2. This aircraft has been passed by the reference aircraft and will not be counted into the queue length of the reference aircraft. In other words, the departure queue of an aircraft is defined as the number of flights that have a takeoff time during its taxi-out and the arrival queue is defined as the number of flights that have a gate in time falling into its taxi-in duration.

The difference in the calculation of queue lengths from the previous two papers is illustrated by an example below.

TABLE 1. Illustration of different queue length calculation methods

Carrier	Gate-out	Wheels-off	Dep_Queue (Kondo)	Dep_queue (Idris)
USA	6:57:00	7:13:00		
NKS	7:00:00	7:15:00		
NWA	7:00:00	7:18:00		
UAL	7:02:00	7:19:00		
UAL	7:04:00	7:22:00		
DAL	7:08:00	7:29:00		
FLG	7:08:00	7:26:00		
NWA	7:10:00	7:24:00	**7**	**5**
AAL	7:14:00	7:27:00		

According to the definition by the APO model the departure queue for NWA at 7:10 am is seven, which is the number of aircraft on the airport surface at its gate out time. The departure queue for that flight is five according to Idris et al.'s definition because it has passed the two flights DAL and FLG that had a gate out time of 7:08 am but took off later than the NWA flight.

The queuing model proposed by Idris et al for taxi out estimation assumed takeoff queue to be the primary factor affecting the taxi out time of an aircraft. They set up different combinations of carriers and runway configurations as subsets. The data of the case study that they presented in the paper contained a total of 56 subsets. The downstream restrictions were not considered as separate variables but were assumed to be a part of the departure queue. Idris et al stated that, aircraft that experienced long taxi out times due to passing and restrictions would have long take off queues. For all the subsets, a probability distribution function (PDF) is developed that gives the probability of a queue forming depending on the number of aircraft present on the airport surface at that particular time. An average taxi out time is calculated over all possible queue sizes and then a second-order equation is fitted to these values. Their model was compared to the running average model that is used in the ETMS and showed a reduced mean absolute error. The model predicted 66% of taxi out times within 5 minutes of actual time and is applicable when the number of aircraft present on the airport surface is known.

The Enhanced Traffic Management System (ETMS) model(9) estimates the taxi-out time using the running averages of the last two weeks. The limitation of this model is it does not take into consideration the important factors affecting the taxi-out time of an aircraft such as runway configuration. Shumsky(10) proposed two linear models to predict the taxi-out time of an aircraft. One was a static model and the other was a dynamic one. The static model uses the variables such as carrier, runway configuration, weather and a measure of airport congestion. To explain airport congestion Shumsky projected two different measurements, the number of pushbacks in a given time period around the pushback of the aircraft, and the number of departing aircrafts present on the runway at the pushback time. The results of this study showed that estimations using the queue size were better than using the number of aircrafts on the runway as a measurement for airport congestion. Shumsky also claimed that the static model was as good as the dynamic model for short time horizon, such as 15 minute period. Nevertheless, for longer time horizon the static model yields superior results.

III. PROPOSED REGRESSION MODEL

This study proposes a set of linear equations to model the taxi-in and taxi-out times. Explanatory variables include arrival and departure queue lengths, runway configuration, arrival and departure runways, and dummy variables indicating time of day and Expect Departure Clearance Time (EDCT) that reflects air traffic flow management activities. Arrival and departure queue lengths and runway configuration have been discussed extensively in the literature review and were widely accepted as major causal factors of taxi-in and taxi-out delay. The information of arrival and departure runways in use are also important because it gives the distance from gates to the end of the runway and the distance from runway exits to gates. Peak and non-peak hours in the day could cause contrasting performance of taxi-in and taxi-out delay due to different gate constraints. In addition, flights experiencing EDCT could perform different from others. Dummy variables are set up for the time of the day and EDCT to account for these effects. Considering the physical interaction between aircraft in the taxiway systems, quadratic terms of the queue lengths are introduced in this regression model. Similar as the APO model, flights are grouped according to carriers and seasons and the flights with taxi times in the upper 25 percent are filtered from the data set as outliers. The case study of this model with 2007 data at LGA shows a higher R square value when compared to other existing models.

A. *Explaination of Variables*

Departure and arrival queues are calculated following the method proposed in Idris et al's paper which has been described in detail in the literature review.

Expected departure clearance times (EDCT)
The traffic management personnel assesses the imbalance of air traffic demand and the capacity of one airport and come up with a plan of holding flights at their origin airports by assigning them expected departure clearance times. Once the EDCT time is allotted the flights have around 15 minutes to depart, otherwise, they will be assigned a new EDCT time which means more schedule delays. Dummy variable is set to be 1 if one flight experienced EDCT or 0 if it did not.

Time of the day
Peak and non-peak hours have different gate constraints which per se affect taxi times. After scrutinizing the scheduling, we divide a day at a specific airport into various time windows. For instance, at LGA, we define the different time of the day into four windows, from 6:00am to 9:00am, from 9:00am to 2:00pm, from 2:00pm to 9:00pm, and after 9:00pm. For each time window, dummy variable is set to be 1 if one flight falls into that window or 0 if it did not.

Runway configuration
For each runway configuration, the dummy variable is set to be 1 if the configuration was operated while one flight taxiing-in or taxiing-out or 0 if it was not.

Arrival and departure runways in use
Arrival and departure runways in use define the distances from gates to the end of runway(s) and the distances from runway exist(s) to the gates. Nevertheless, this information is hard to obtain. In ASPM data that we used to conduct the case study, there are no arrival and departure runways in use recorded. Fortunately, we can find some airports, LGA as one of them, which has only one arrival runway and only one departure runway. Thus, given the runway configuration, it is

easy to know the arrival and departure runways in use. For modeling other airports with more complex runway configuration, additional database, such as Performance Data Analysis and Reporting System (PDARS), need to be used for obtaining such information.

B. Case Study and Data Sources

Airports with longest taxi-out times are typically those with higher volume of air traffic. These airports are mostly either hub airports or focus cities for airlines. According to BTS, for 2007, the top three in the list of airports with longest ground times waiting for takeoff in 2007 were from the New York area and LGA was ranked at number three with average taxi-out times of 29 minutes. As we have described in the Section 3.1., not only the runway configuration but also the information of specific runways that flights are assigned to will affect the taxi times. Among the three New York airports, LGA is an ideal airport for our case study because it has only two cross runways, one for arrival and one for departure. The data for the case study was downloaded from aviation system performance metric (ASPM complete).

C. Regression Results and Comparision

With the same data, we conducted the regressions of our proposed model and the existing model used to calculate the nominal taxi times recorded in ASPM database. The comparison of the performance of the two models is shown in table 2. The proposed model has an average R^2 value of 0.758 for taxi-out estimation across all groups while the average R^2 value of the other model is 0.429. In addition, the standard error of the R^2 values for the proposed model is smaller.

TABLE 2. Comparison of the performance of different taxi time regression models

R-square Statistics	***Alternate model***	***Existing model***
Mean	0.758	0.429
Standard Error	0.004	0.008
Median	0.753	0.434
Mode	0.738	0.455
Standard Deviation	0.044	0.084
Sample Variance	0.002	0.007
Kurtosis	0.814	-0.120
Skewness	0.627	-0.303

IV. PREDICTING TAXI TIMES

A. Iterative Algorithm

Given the regression results and other inputs from flight scheduling, we propose an iterative algorithm to predict the taxi out times. The basic idea is to revise arrival and departure queue lengths and update the taxi-out times of the flights in each iteration until the difference between two iterations becomes less than the convergence parameter set up at the beginning. The pseudo code of the algorithm is as below in fig. 2. Initially the arrival and departure queue lengths are set as zero. The iteration count variable n is set as one and convergence parameter is defined as 0.005. Given the estimated coefficients and other input variables, the taxi-in time and taxi-out times can be calculated. Given gate out times and arrival times, we can calculate departure times for departure flights and gate in times for arrival flights. Assuming there are no gate constraints holding arrival flights from getting a gate, we only check the extra taxi-out times that could cause by departure capacity. The 15-minute airport departure rate (ADR) is used as departure capacity of the airport. With the previous calculation, we can check if the 15-minute ADR is exceeded or not. If so, affected flights are postponed to next 15-minute time window. The same procedure is repeated until no demand exceeds supply in all 15-minute time windows in the day. Assuming there is no over passing, we can calculate the arrival or departure queue lengths and then the taxi-in or taxi-out time for each flight. Compare the two sets of taxi-in and taxi-out times mentioned so far, if the differences are smaller than the convergence parameter, the iterative algorithm stops, otherwise, the iteration counts increase one unit and the iteration continues from calculating the departure times for departure flights and gate in times for arrival flights.

1. Initialization queue length: $x_o^{(0)} \leftarrow 0$ and $x_i^{(0)} \leftarrow 0$, iteration count $n \leftarrow 1$, convergence parameter $\varepsilon = 0.005$

2. Given estimated coefficients from regression model, calculate $t_o^{(n)}$ and $t_i^{(n)}$

3. Given gate out time g_o and arrival time a_i, calculate departure time $d^{(n)} = g_o + t_o^{(n)}$ and gate in time $g_i^{(n)} = a + t_i^{(n)}$

4. Check 15-minute total departures. If the capacity (ADR) is exceeded, affected flights are moved to the next time window.

   Stops when all ADR constraints are satisfied.

5. Calculate departure and arrival queue lengths $x_o^{(n)}$ and $x_i^{(n)}$, assuming no overpassing

6. Given estimated coefficients from regression model, calculate $t_o^{(n+1)}$ and $t_i^{(n+1)}$

7. Convergence test: If $\sum_{f_a}\left(t_i^{(n+1)} - t_i^{(n)}\right)/F_a < \varepsilon$ and $\sum_{f_d}\left(t_o^{(n+1)} - t_o^{(n)}\right)/F_d < \varepsilon$, stop, else $n \leftarrow n+1$ and go to step 3

FIGURE. 2. Pseudo Code of the Alternative Algorithm

B. Case Study and Performance of the Algorithm

We picked one day in 2007, July 13th, at LGA to test the performance of the algorithm. More experiments should be conducted later to get a more general idea about the performance. It shows that the model is able to predict 74% of taxi-out times within five minutes of the actual times. With a different date set, the model proposed by Idris et al predicted 66% of taxi-out times within five minutes of actual times. Table 3 list the descriptive statistics when comparing the predicted taxi-out times and actual taxi-out times recorded in

ASPS data. Fig. 3 demonstrates the comparison of average taxi-out times for different hours of the day. It is observed that in the evening, there are larger discrepancies between predicted taxi-out times and actual taxi-out times. It could be caused by the gate constraints that we have ignored in our iterative algorithm or other factors. To predict taxi times more accurately, it is worth of more investigation by looking into surface movement data, observing the real-time operations at airports, and evaluating the impact of gate constraints on arrival queues.

TABLE 3. Comparison of Predicted Taxi-out times and Actual Taxi-out times

Statistics	***ACTTO***	***CALTO***
Mean	18.55	18.95
Standard Error	0.23	0.22
Median	18.00	18.38
Mode	12.00	19.32
Standard Deviation	5.48	5.25
Sample Variance	29.98	27.56
Kurtosis	0.00	0.72
Skewness	0.63	0.47

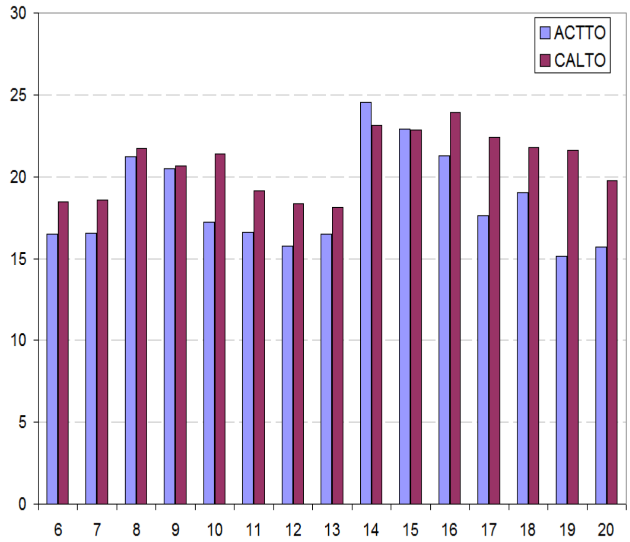

FIGURE 3. Comparison of Taxi-out times in different times of the day

V. CONCLUSION

This paper proposed a set of regression equations to model the taxi times at airports by considering the queuing effect, runway configuration and runways in use, EDCT effect, time of day and others. The comparison of the proposed model and the model used to calculate the nominal times recorded in ASPM database show that with the expansion of independent variables, the proposed model explains double of the variation of the taxi times. The paper then presented an iterative algorithm for predicting taxi times. The inputs for the algorithm include the estimated coefficients from aforementioned regression model, flight gate out times or arrival times. ADR is taken as the airport departure capacity. Procedures are taken to ensure the departure capacity is not exceeded in each iteration. The algorithm is tested with one day operations in 2007 at LGA. The predicted results are compared with the actual taxi out times recorded in ASPM. Overall, 74% of predicted value falls into the range within five minutes of the actual times. This is higher than the 66% claimed by one of the existing model, although with data from a different airport.

REFERENCES

[1] . Idris, J.P. Clarke, R. Bhuva and L. Kang. “Queuing Model for Taxi-Out Time Estimation”. Air Traffic Control Quarterly, September 11th 2001

[2] B. Goldberg and D. Chesser. ”Sitting on the Runway: Current Aircraft Taxi Times Now Exceed Pre-9/11 Experience”. Bureau of Transportation Statistics Special Report, May 2008.

[3] N.Pujet, B. Delcaire, E. Feron. “Input-Output Modeling and Control of the Departure Process”. Proceedings of AIAA Guidance Navigation and Control Conference, 1999.

[4] M.O. Ball, R.L. Hoffman, D. Knorr, J. Wetherly and M. Wambsganss. “Assessing the benefits of Collaborative Decision Making in Air Traffic Management”. Air Management R&D Seminar, June 2000.

[5] A. Kondo. “Derivation of Unimpeded Taxi Times”. Office of Aviation Policy and Plans. Federal Aviation Administration.

[6] Website of FAA program and initiatives. http://www.faa.gov/about/initiatives/nextgen

[7] E. Mueller and G.B. Chatterji. “Analysis of Aircraft Arrival and Departure Delay Characteristics”. Proceedings of AIAA Aircraft Technology, Integration, and Operations Forum at Aviation Week’s Aerospace Expo, Los Angeles, CA, October 2002.

[8] ETMS, Enhanced Traffic Management System, Functional Description, Report No. VNTSC-DTS56_TMS-002, Volpe National Transportation Systems Center, U.S. Department of Transportation, Cambridge, MA, 2000.

[9] R.A. Shumsky. “Dynamic and Statistical Models for the Prediction of Aircraft Take-Off Times”. Ph.D. Thesis Operations Research Center, MIT, Cambridge, MA. 1995.

[10] J.P. Heimlich. “Commercial Jet Fuel Supply: Impact on U.S. Airlines”. Aviation Subcommittee of the committee on Transportation and Infrastructure of the House of Representatives, February 2006.

[11] Website of the European Organization for the Safety of Air Navigation. http://www.eurocontrol.int