Prediction of Pushback Times and Ramp Taxi Times for Departures at Charlotte Airport

Hanbong Lee[1], William J. Coupe[2], and Yoon C. Jung[3]
NASA Ames Research Center, Moffett Field, CA 94035, USA

When optimizing the takeoff sequence and schedule for departures at busy airports, it is important to accurately predict the taxi times from gate to runway because those are used to calculate the earliest possible takeoff times. Several airports like Charlotte Douglas International Airport show relatively long taxi times inside the ramp area with large variations, with respect to the travel times in the airport movement area. Also, the pushback process times have not been accurately modeled so far mainly due to the lack of accurate data. The recent deployment of the integrated arrival, departure, and surface traffic management system at Charlotte airport by NASA enables more accurate flight data in the airport surface operations to be obtained. Taking advantage of this system, actual pushback times and ramp taxi times from historical flight data at this airport are analyzed. Based on the analysis, a simple, data-driven prediction model is introduced for estimating pushback times and ramp transit times of individual departure flights. To evaluate the performance of this prediction model, several machine learning techniques are also applied to the same dataset. The prediction results show that the data-driven prediction model is as good as the machine learning algorithms when comparing various prediction performance metrics.

I. Introduction

NASA has developed many advanced air traffic management capabilities to integrate arrival, departure, and surface operations under the Airspace Technology Demonstration 2 (ATD-2) project.[1-3] One of the main capabilities in this project is the Surface Predictive Engine which generates an accurate surface model and associated schedule of individual flights. The Surface Predictive Engine consists of the surface modeler and the surface scheduler.[4] The surface modeler integrates the data feeds and inputs from operational users like airlines, tower and ramp controllers, as well as System Wide Information Management (SWIM) feeds from the FAA, and estimates the unimpeded taxi times from gate to runway based on the current aircraft locations and the expected trajectories on the airport surface. These predicted taxi times are used to calculate the earliest possible takeoff times of departures. The surface scheduler generates the best runway schedule based on the estimated runway arrival times of departures and arrivals, subject to various operational constraints.[5] For departures, therefore, it is important to predict the taxi-out times from their gates to the assigned runway accurately, while scheduling takeoff times to maintain an optimal runway throughput and improve efficiency in airport operations.

Various approaches to predicting taxi times of departures have been proposed and investigated, including queuing models, discrete event simulations, linear regressions, and other machine learning techniques. Queuing models consider a runway and a departure queue near the runway as a server and a single queue, respectively, and obtain the taxi time as the sum of the unimpeded taxi time from gate to runway and the wait time in the queue. These models have been improved by considering the effect of taxiway congestion and expanding them to a queuing network model to cover multiple bottleneck areas on the surface.[6-8] Using more accurate taxi time data like the Airport Surface Detection Equipment, Model X (ASDE-X) surface surveillance data, linear regression models were also developed to predict departure taxi times at several major airports in the United States.[9-12] Similarly, various regression methods such as multiple linear regression, least median squared linear regression, support vector regression, model trees, and

[1] Aerospace Engineer, NASA Ames Research Center, Mail Stop 210-6, Moffett Field, CA 94035, hanbong.lee@nasa.gov. AIAA Member.
[2] Aerospace Engineer, NASA Ames Research Center, Mail Stop 210-6, Moffett Field, CA 94035, william.j.coupe@nasa.gov. AIAA Member.
[3] Aerospace Engineer, NASA Ames Research Center, Mail Stop 210-6, Moffett Field, CA 94035, yoon.c.jung@nasa.gov, AIAA Senior Member.

fuzzy rule-based systems, were applied to several airports in Europe for taxi time prediction problems.[13,14] Machine learning algorithms were also introduced to predict the total taxi-out time from gate to runway at major US airports, considering several factors affecting the taxi times. The machine learning techniques applied to solve the taxi time prediction problem included neural networks, linear regression models, support vector machines, k-nearest neighbors, and random forest.[15-17] Some of them were compared with the simulated taxi times from a fast-time simulation model, which can be an alternative taxi time prediction approach, to evaluate the prediction performance.[16]

Most taxi time prediction approaches listed above focused on the total taxi-out times from gate to runway. Depending on the location of aircraft taxiing on the ground and its control authority, however, the taxi-out time can be broken down into two parts: taxi times in the ramp area and the airport movement area. In addition, the ramp taxi time of a departing aircraft is divided into the pushback time when being pushed backward by a tug and the pure ramp transit time when moving forward under its own power. In other words, the total taxi-out time, T_{out}, can be expressed as the sum of the pushback process time, PbT, the pure transit time from a designated area for engine spooling up near gate to spot in the ramp area, $T_{g\rightarrow s}$, and the transit time from spot to runway in the airport movement area, $T_{s\rightarrow r}$.

$$T_{out} = PbT + T_{g\rightarrow s} + T_{s\rightarrow r} \tag{1}$$

Depending on the airport layout and geometry, ramp taxi times, as well as their variations, can be as large as those in the airport movement area. For example, Charlotte Douglas International Airport (CLT), one of the busiest airports in the US, shows relatively long taxi times in the ramp area for both departures and arrivals, with large variations.[8] Because of its unique layout of concourses, westbound departures from Concourse E usually have long taxi distance to spots, resulting in high taxi time values in the ramp area.

In addition, the pushback procedure has not been considered in taxi time estimation appropriately. Estimating the duration of the pushback procedure is quite challenging due to a lot of uncertainties involved in the process. The pushback process time can be defined as the duration from the off-block time after obtaining pushback clearance to the moment when the departing aircraft starts taxiing under its own power. During the pushback process, many events occur both in sequential and parallel ways. These events include at least, but are not limited to, aircraft pushback by a tug, engines spooling up, tug detachment, communication between pilot and ground crew, and communication between pilot and ramp controller for taxi clearance. Each event itself has some uncertainties in time for completion. Therefore, the entire pushback process time, which is the sum of durations of all sequential steps, has large variations flight by flight, making it difficult to estimate the finish time of the pushback process for a departure flight.

Previous work tried to model the various processes within the ramp area, but it faced the challenge that accurate data for the different events were not available at that time.[18] Fortunately, the detailed ramp operations data with the acceptable level of accuracy are now obtainable for CLT due to the deployment of the ATD-2 system and the active participation of ramp tower controllers in the ATD-2 project, which makes it possible to develop separate prediction models for pushback times and ramp taxi times.

The main contributions of this paper include the following: First, detailed pushback process time and ramp taxi time analyses are performed based on more accurate data leveraged from the ATD-2 system deployed in the CLT ramp tower. Second, a decision tree for pushback process time is developed as a data-driven prediction model for determining the pushback times of departing aircraft. Next, a simple prediction model for ramp taxi times of departures that can be used in real time operations is developed by focusing on the pure transit time in the ramp area from the designated area for engine spooling to the assigned spot, based on the historical data and standard taxi routes. Lastly, machine learning techniques are applied to evaluate the prediction accuracy of the proposed prediction models and investigate the possibilities of prediction performance improvements.

This paper is organized as follows. In Section II, actual pushback times and ramp taxi times of departures operated at CLT are analyzed to identify the characteristics of the surface traffic at this busy airport. A simple prediction model for pushback time estimation is introduced in Section III, while comparing its prediction accuracy with the one using machine learning techniques. In Section IV, a ramp taxi time prediction model based on historical data is proposed and evaluated with machine learning methods. Lastly, Section V concludes with future work.

II. Pushback Time and Ramp Taxi Time Data Analysis

A. Data inputs

In late 2017, the Phase 1 Baseline Integrated Arrival, Departure, Surface (IADS) traffic management capability of NASA's ATD-2 system was deployed both in the air traffic control tower and ramp control tower at Charlotte airport.[4] Through this system, NASA has been collecting all the operations data for the airport. Among the collected data, the inputs to the system made from ramp controllers such as pushback approval time, pushback hold, and taxi clearance time can be used to obtain the accurate pushback process times of departures. Actual off-block times and spot crossing times from various data sources like ramp controller input and surface surveillance radar are used to compute the taxi times of departures in the ramp area.

B. Pushback Time Variation Analysis

Figure 1 shows the histogram of actual pushback times at CLT, collected for one month in August 2018. For 20,595 departures during this period, the average value and standard deviation of the pushback times are 4.72 minutes and 2.27 minutes, respectively. In this analysis, departures having invalid values or missing inputs are excluded. If neither pushback approval nor taxi approval times are available, the flight is excluded because its pushback time cannot be calculated. If the state of a flight is suspended, returned to gate, unknown, or canceled, the flight is also excluded in this analysis. Some outliers like less than 30 seconds or greater than one hour in pushback time are also removed. Despite this data filtering, it is noted that some data may have incorrect time values due to the errors in input or measurement.

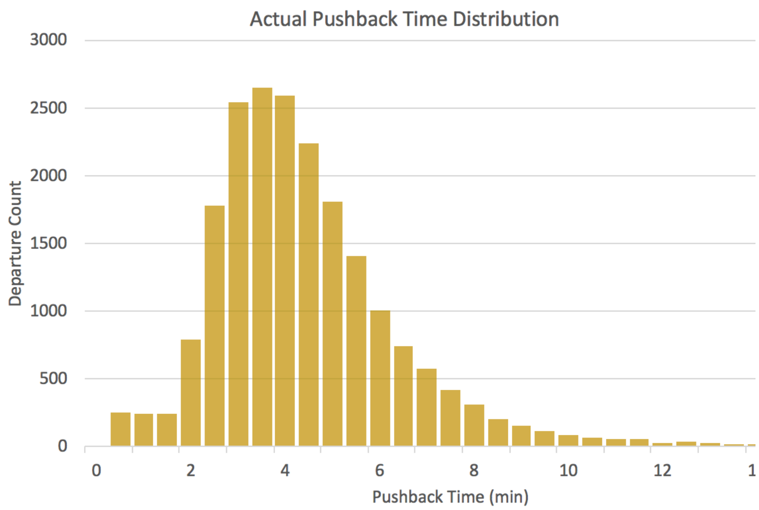

Figure 1. Actual pushback time distribution at CLT

The pushback time statistics show some variations by several factors affecting the pushback processes such as gate location, aircraft type, and carrier. First, gate location (or segmented ramp area grouping adjacent gates) may lead to different pushback durations due to the geometric characteristics around the gate and boarding/loading equipment capabilities assigned to the gate. For example, a departing aircraft from the gate inside the concourse B-C alley may have to wait until the other aircraft sharing the same pushback spot for engine spooling is cleared. For this reason, the average pushback time of departures inside the B-C alley (4.53min) is higher than the pushback times from the tip area at concourse B and C (3.70min). Also, different aircraft types can result in different pushback process times. Usually, heavy aircraft take more time to complete the required pushback processes. Some gates at concourse D that are frequently used for international flights show a noticeably longer pushback time (6.17min) compared to other gates. Table 1 shows the average pushback times by wake turbulence category, supporting this inference. Figure 2 shows these pushback time variations depending on ramp area and aircraft type in detail. Furthermore, each airline has its own pushback procedure and guidelines, which may produce differences in pushback times, even for the flights having the same aircraft type at the same gate. In this paper, however, the pushback times by carrier are not shown due to the sensitivity of those data.

Table 1 Average Pushback Times by Wake Turbulence Category

Wake Turbulence Category	B	D	E	F
Average Pushback Time (in minutes)	5.30	4.67	4.50	1.67

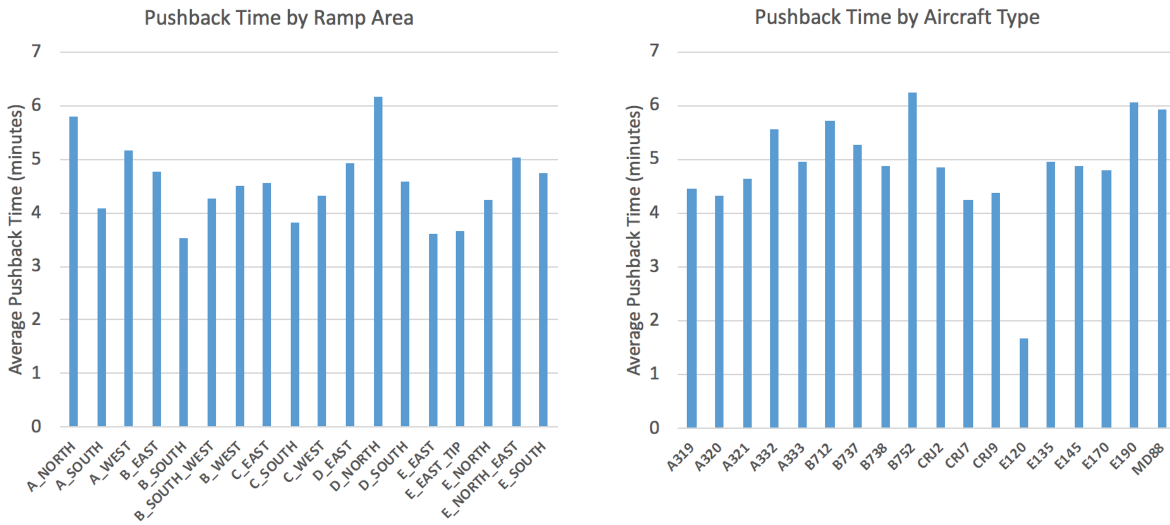

Figure 2. Pushback time variations by ramp area (left) and aircraft type (right)

C. Ramp Taxi Time Variation Analysis

Similarly, the histogram of actual ramp transit times at CLT in August 2018 is illustrated in Figure 3. The ramp transit time from the designated area for engine spooling to the assigned spot is obtained by calculating the difference between the spot crossing time and taxi clearance time. The spot crossing time, or the airport movement area entry time, of a flight is obtained from the ASDE-X surface surveillance data. The taxi start time of the flight is assumed to be the same as the taxi clearance time, which is available from the manual input made by the ramp controller through the Ramp Traffic Console (RTC), the user interface tool for ramp controllers provided by the ATD-2 system. The delays from taxi clearance to the actual taxi start, such as the communication delay between ramp controller's direction and pilot's action, are ignored in this calculation.

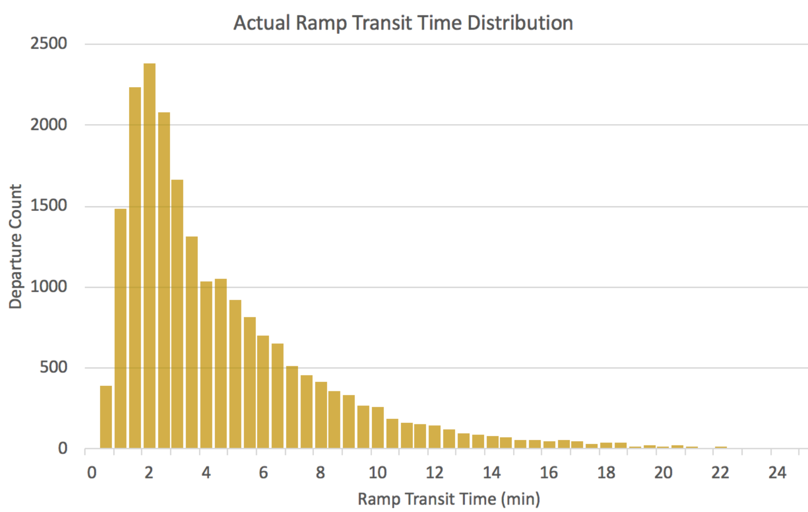

Figure 3. Actual ramp transit time distribution at CLT

The average value and standard deviation of the pure ramp transit times for 21,093 departures having valid time values are 4.93 minutes and 4.39 minutes, respectively. It is noted that the variation of the ramp transit times is somewhat large. These large variations come from the unique terminal layout at this airport shown in Figure 4, causing various taxi distances from gate to spot. The taxi distance from gates in Concourse E to Spot 29 (a spot near the concourse for Eastbound flights), for example, can be short, whereas the distance from the same gates to Spot 9 (a spot near the opposite side of the main terminal building for Westbound flights) is very long. This observation indicates that the primary contributor to ramp taxi times at CLT is the taxi distance from gate to spot.

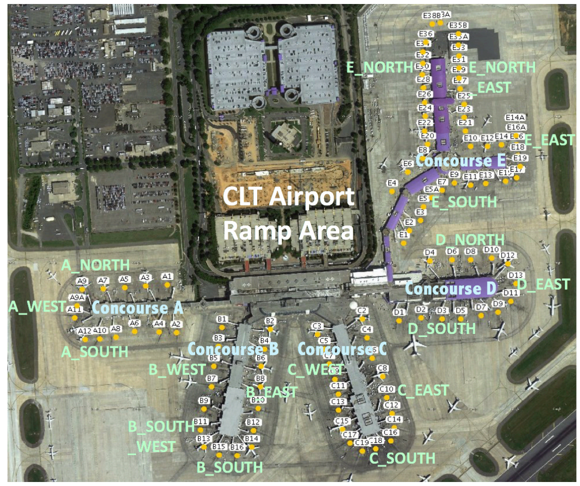

Figure 4. CLT ramp area diagram

III. Pushback Time Prediction

In this section, a simple prediction model for pushback time estimation based on two main criteria is introduced. This model is applied to the actual data to evaluate the prediction performance. Then, several machine learning techniques are developed and tested with the same dataset to compare their prediction accuracies.

A. Data-driven Pushback Time Prediction Model

From the actual pushback time analysis in the previous section, it is found that the pushback times are dependent upon gate location and aircraft type. Several neighboring gates in the same concourse are grouped and called the 'ramp area' in the ATD-2 system, and adopted in this paper. According to the concourse and gate location, CLT has several ramp area elements per concourse, e.g., A_NORTH, A_WEST, and A_SOUTH for A concourse. Figure 4 shows these ramp area names with the CLT terminal layout. The two main criteria, ramp area and aircraft type, are used to create a decision tree for pushback times. Based on the statistics of historical data, a likely pushback time value for the flights belonging to each specific ramp area and aircraft type is defined by taking a median value from historical data. By determining all the likely values available for the ramp area and aircraft type branches, the decision tree for pushback times is structured, as shown in Figure 5. If the sample size for a branch is small or the likely values are not available, the default value, the median value from the entire set of data samples (4.33min), will be used. Then, this decision tree is used to determine the estimated pushback time of a departure, given the assigned gate and the aircraft model.

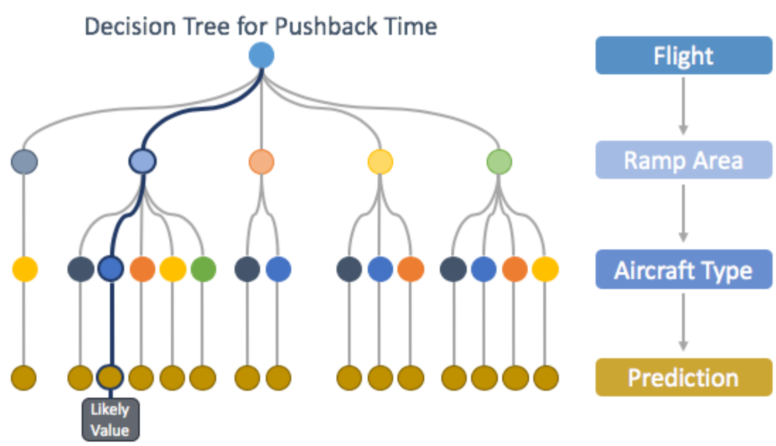

Figure 5. Decision tree structure for pushback time

As a reference for comparison, we can consider a simpler model using a single, default value that can be applied to all departures, regardless of ramp area and aircraft type. To minimize the deviations, the default value is set as the median of all the pushback time data, 4.33 minutes.

Figure 6 shows the pushback time prediction accuracy from the decision tree model and the reference model. Table 2 shows the statistics of the pushback time prediction accuracy from the two prediction models, including the mean and median values of pushback time difference (actual – predicted), mean absolute error, root mean square error, and the percentages of flights within ±1, 3, and 5-minute error. The decision tree model performs better by taking the ramp area and aircraft model of individual departures into account, but the gain is not significant.

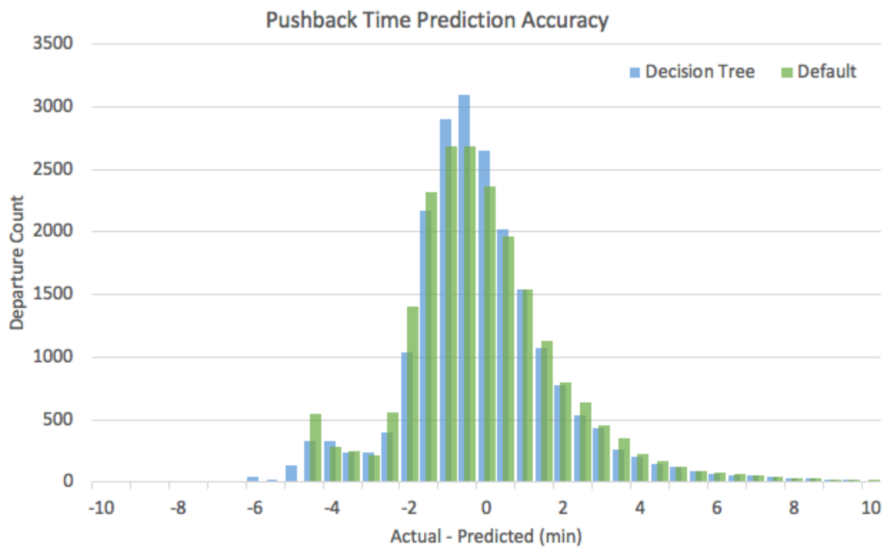

Figure 6. Pushback time prediction accuracy

Table 2 Pushback Time Prediction Accuracy Statistics from Two Models

Prediction Model	Decision Tree	Default value
Average of errors (min)	0.37	0.39
Median of errors (min)	0.00	0.00
Mean Absolute Error, MAE (min)	1.40	1.49
Root Mean Square Error, RMSE (min)	2.24	2.31
Flights within ±1min error	52.1%	47.3%
Flights within ±3min error	89.8%	89.0%
Flights within ±5min error	96.9%	96.8%

B. Machine Learning Algorithms

In this paper, we try six different machine learning algorithms for the pushback time prediction. These algorithms include Linear Regression (LR), Support Vector Regression (SVR), Lasso linear regression (Lasso), *k*-Nearest Neighbors (*k*NN), Random Forest (RF), and Neural Networks (NN). Linear Regression (LR) finds a linear relationship between a scalar dependent variable y and one or more explanatory variables denoted X. For prediction, it can be used to fit a predictive model to an observed dataset of y and X values, with coefficient w to minimize the residual sum of squares between the observed variables and the responses predicted by the linear approximation. After developing such a model, if an additional value of X is given without its accompanying value of y, the fitted model can make a prediction of the y value. Support Vector Machines (SVM) are supervised learning models with associated machine learning algorithms that analyze data and recognize patterns. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. Support Vector Regression (SVR) is an extended version of SVM to solve regression problems. The Lasso regression, which stands for least absolute shrinkage and selection operator, is another linear regression model having a different objective function for data fitting. This linear model selects a subset of the provided covariates for use in the final model for the better prediction performance. *k*-Nearest Neighbors (*k*NN) algorithm is a non-parametric method used for classification and regression. The input consists of the *k*-th closest training examples in the feature space. In *k*NN regression, the output is the property value for the object. This value is the average of the values of its *k* nearest neighbors. Random Forest (RF) is an ensemble learning method for classification and regression that operates by constructing a multitude of decision trees at training time and computing the class that is the mode of the classes output by individual trees. These six algorithms will be used to estimate the pushback times of departures in the test dataset and evaluate the prediction performance of the simple model based on the Decision Tree (DT). These machine learning algorithms are implemented in Python and run using the *scikit-learn* library.[19]

For the variables determining the pushback duration of a departing aircraft, there exist several features that can contribute to the pushback process duration among the available data. Those features include ramp area (gate groups accommodating adjacent gates), aircraft type, carrier, actual pushback time of a day in hour, gate conflict, and the restriction of Expected Departure Clearance Time (EDCT) and Approval Request (APREQ). Among these features, ramp area and aircraft type were already used to define the decision tree for the pushback time prediction. In addition, aircraft carrier is considered as a feature because each airline can have a different procedure and guideline for pushback processes. It may take a longer time for a departing aircraft to push back when traffic demand is high. Considering that CLT has a peak and valley traffic pattern on a daily basis, the actual pushback time of the day is added to the feature set. Gate conflict situations where a departure is still occupying a gate, but a landing arrival is assigned to the same gate, can affect the pushback time as well. According to the one-month data, although the average pushback time is almost the same, the standard deviation for the departures subject to potential gate conflict (4.38min) is higher than normal departure's (2.22min). Lastly, the existence of Traffic Management Initiative (TMI) restrictions like Expected Departure Clearance Time (EDCT) and Approval Request (APREQ) are also considered as features because these controlled flights need more attention from the controllers in their ground movement to meet the given controlled takeoff times. The actual data show that the APREQ flights have relatively higher pushback times on average (5.30min) with larger variations, compared to the other departures (4.67min) and that the EDCT flights also have slightly higher average values (5.16min) with larger variations than others (4.70min).

From the given dataset, a total of 67 features were defined and commonly used in running the machine learning algorithms. These features include 18 binary variables for ramp area (3 for A concourse, 4 for B concourse, 3 for C concourse, 3 for D concourse, and 5 for E concourse), 23 binary variables for air carriers, 23 binary variables for aircraft types, one binary variable for a potential gate conflict with an arrival, one variable for the pushback time of a day (in hour), and 2 binary variables for TMI (one for APREQ and the other for EDCT).

The same one-month data in August 2018 was used to run the machine learning algorithms and compare their prediction performances with the result from the decision tree (DT) model. The two-thirds of the data were used as a training dataset for the machine learning algorithms, whereas the remaining one-thirds was used as a test dataset. The whole one-month data were rotated three times in training and test partition so that each day had a test result to be compared to the actual data. For example, the prediction models were trained with the departure flight data from August 1, 2018 to August 20, 2018, and then used to obtain the predicted values of pushback times for the flights between August 21, 2018 and August 31, 2018. In a similar way, the flight data from August 11 to August 20 was set as the next test dataset with the training dataset from the remaining flight data. Lastly, the flight data between August 1 and August 10 was tested for the pushback time prediction after trained with the flight data between August 11 and August 31. The resultant prediction values of the pushback times for the flights in August 2018 were compared with the actual pushback times in order to calculate the prediction accuracy of each prediction model.

The pushback time prediction results from different machine learning algorithms are shown in Table 3. The prediction accuracy was evaluated with various performance metrics in statistics, including the mean and median values of the prediction errors (actual minus predicted values), mean absolute error, and root mean square error, and the percentages of flights within the given accuracy window like ±1, ±3, and ±5 minutes. The prediction performance metrics of the decision tree model (DT) were also added to the comparison table for evaluation.

Table 3 Pushback Time Prediction Accuracy Comparison with Various Machine Learning Models

Prediction Model	LR	SVR	Lasso	*k*NN	RF	NN	DT
Average of errors (min)	0.00	0.39	-0.01	0.05	0.00	-0.01	0.37
Median of errors (min)	-0.38	0.00	-0.40	-0.20	-0.31	-0.38	0.00
Mean Absolute Error (min)	1.42	1.45	1.46	1.56	1.47	1.43	1.40
Root Mean Square Error (min)	2.19	2.28	2.22	2.37	2.25	2.20	2.24
Flights within ±1min error	47.9%	49.4%	45.4%	45.8%	47.9%	47.4%	52.1%
Flights within ±3min error	90.6%	89.1%	90.4%	87.7%	89.3%	90.5%	89.8%
Flights within ±5min error	97.4%	96.8%	97.3%	96.6%	97.2%	97.4%	96.9%

According to the metrics in Table 3, the prediction accuracy has no significant differences between machine learning algorithms, although the *k*NN model shows a little lower accuracy. Also, it seems that the SVR model tries to make the median value of the errors close to zero, instead of the average value. The mean and median values of the prediction errors are almost zero or less than 25 seconds. The MAE and RMSE values suggest that the deviations of the predicted pushback times from the actual ones are around 90 seconds regardless of methods, but it is expected that about half of departures complete their pushback processes within ±1 minute from the predicted time. The data driven prediction model using a decision tree (DT) in the last column of the table also shows the same level of prediction accuracy as the machine learning algorithms' accuracies. This result implicates that the decision tree model using only two criteria, ramp area and aircraft type, works well in the pushback time prediction for most departures. The box plot in Figure 7 provides a better visualization for comparison between the different models. In the box plot, the bottom and top lines of the box represent the 1st and 3rd quartiles of the prediction errors, respectively. The middle line of the box shows the median value, whereas the x mark inside the box means the average value. The whiskers (vertical lines) extend from the ends of the box to the minimum value and maximum value, excluding outliers.

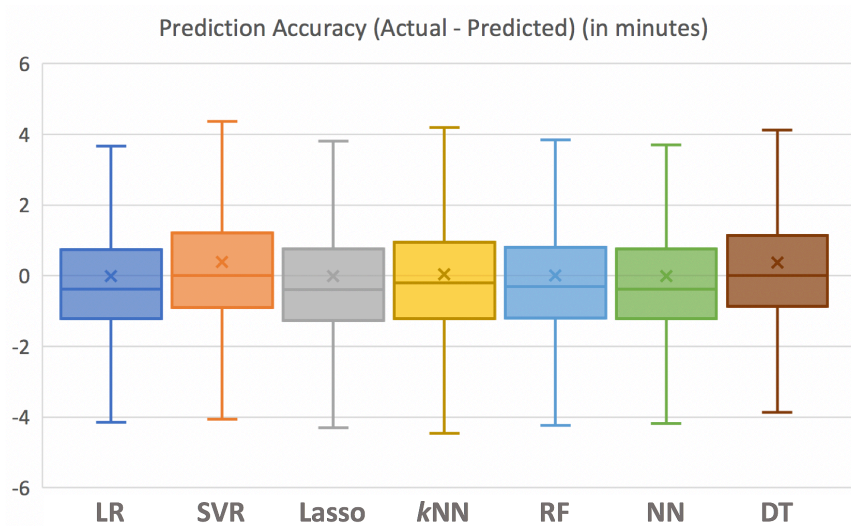

Figure 7. Box plot for pushback time prediction accuracy from various machine learning algorithms

IV. Ramp Taxi Time Prediction

In this section, a simple, data-driven model for ramp taxi time prediction is introduced. The prediction accuracy of this model is first assessed against the actual flight data samples at CLT, and then further evaluated with the test prediction results from a set of machine learning algorithms.

A. Data-driven Ramp Taxi Time Prediction Model

As described in Section II.C., the actual ramp taxi time of a departing aircraft is calculated from the difference between its spot crossing time and taxi clearance time. Given the assigned gate and spot information, the taxi routes can be defined using a node-link model for CLT airport. From the primary taxi routes inside the ramp area, the taxi distance from gate to spot is available for all the combinations of gates and spots. Using the actual taxi times and primary taxi distances of individual departures, we can easily calculate the average taxi speed while a departing aircraft travels from its gate to the assigned spot.

For a departure flight *i* in a set of departures, *D*, the average ramp taxi speed can be obtained by dividing the ramp taxi distance from gate to spot by the actual ramp transit time.

$$Ramp\_Speed_i = \frac{Ramp\_Distance_{g \to s}}{Ramp\_Transit\_Time_i}, \ \forall i \in D \tag{2}$$

From the historical data collected for the several months since the deployment of the ATD-2 system at CLT, we can obtain the distribution of various ramp taxi speed values for departures, as shown in Figure 8. In this graph, some departures having unreasonable speed values for ramp taxi operations (i.e., greater than 30 knots or less than 1 knot) were excluded. The mean, median, and standard deviation of these ramp taxi speeds are 7.0 knots, 6.6 knots, and 3.6 knots, respectively.

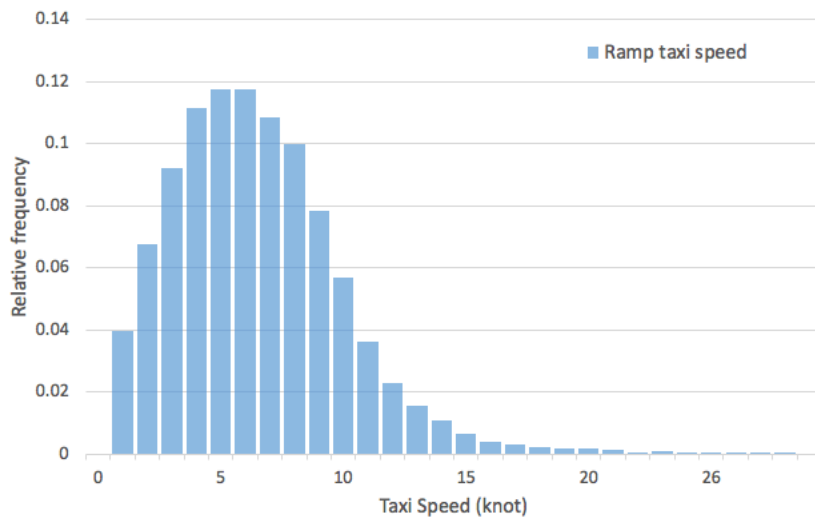

Figure 8. Ramp taxi speed distribution at CLT

There are large variations in the ramp taxi speed distributions due to the uncertainties in ramp operations such as unexpected stopping for safety, waiting before entering the bi-directional taxi lane near Concourse D, and being blocked by other aircraft pushing back. To make it simple to estimate the spot arrival times of departures, it is assumed that all the departures move in the ramp area at a constant taxi speed. This assumption prevents the departures from overtaking each other while taxiing, which can help predict reasonable spot arrival times and sequences for the flights going to the same spot. For this purpose, the median value of 6.6 knots is taken from the statistics above.

Given the assigned gate and spot of a departure flight, the prediction model embedded in the ATD-2 system calculates the ramp taxi distance from gate to spot along the standard taxi route. Then, the model estimates the ramp taxi time of the flight, with the default ramp taxi speed value. In such a manner, the ATD-2 surface modeler operated in the field at CLT can estimate the ramp transit times very quickly. Another merit of this simple prediction model using the mean taxi speed value is that the system can estimate the spot arrival time of the flight who already left its gate, based on the remaining taxi distance from the current position to the assigned spot, in real time.

Figure 9 shows the ramp transit time prediction accuracy of the departure dataset in August 2018 from the data-driven prediction model using the default single taxi speed value and the ramp taxi distances of individual flights. The predicted ramp transit time of each flight is compared with the actual ramp taxi time, and the differences are plotted in the histogram format. Because the median taxi speed value was chosen in the prediction model, the error distribution is almost symmetric at the zero value, but has some deviations at both sides. Table 4 shows the statistics of the ramp taxi time prediction accuracy. The mean and median values of the errors, actual taxi time – predicted taxi time, are both less than 30 seconds, showing a good prediction performance at large. However, the RMSE metric is large because some flights experience a long travel time inside the ramp due to the congestion or other reasons.

Table 4 Ramp Transit Time Prediction Accuracy Statistics

Prediction model	Default taxi speed model
Average of errors (min)	0.07
Median of errors (min)	-0.45
Mean Absolute Error, MAE (min)	2.32
Root Mean Square Error, RMSE (min)	4.00
Flights within ±1min error	37.4%
Flights within ±3min error	76.4%
Flights within ±5min error	90.4%

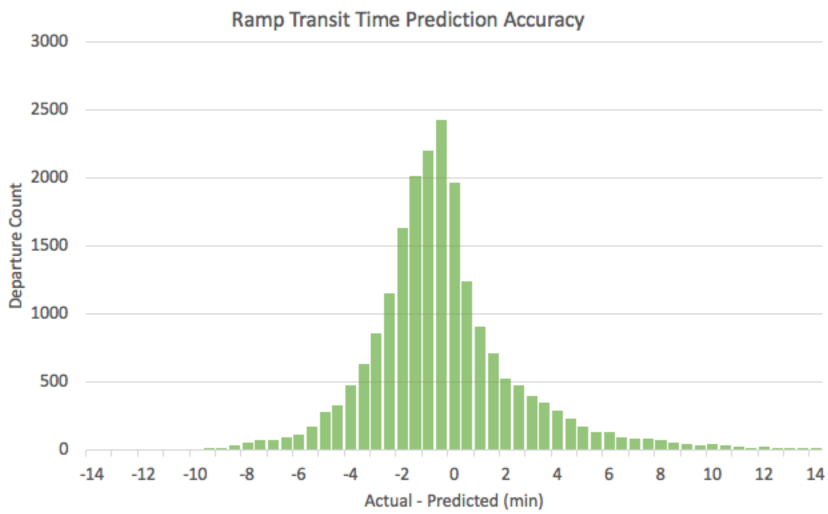

Figure 9. Ramp transit time prediction accuracy

B. Machine Learning Algorithms

In this paper, several machine learning algorithms are applied for ramp transit time predictions. The same six methods listed in Section III.B. are tried in a similar way, but the features used in the prediction models are different to achieve the better prediction performance. Among the flight data elements available from the ATD-2 system, the following features are considered as the variables that could have an impact on the taxi-out time in the ramp area: ramp area (gate groups), spot, aircraft type, carrier, actual pushback time of a day in hour, gate conflict, EDCT and APREQ restrictions, runway configuration, ramp taxi distance from gate to spot, and the number of departures and arrivals taxiing in the ramp.

Ramp area and spot are chosen as the basic features for the machine learning algorithms because they determine the taxi route of a departure taxiing in the ramp area. Aircraft type, carrier and pushback time of a day are also considered for the similar reasons to the pushback time prediction. According to the actual flight data in August 2018, departures having a gate conflict possibility experienced a slightly longer ramp transit time on average (5.06min), compared to the other flights (4.85min). Departures subject to APREQ and EDCT also showed longer ramp transit times by 0.21min and 1.16min, respectively, although they are only 8.9% and 4.4% out of all departures. Based on these observations, gate conflict, APREQ, and EDCT are included as features to improve the ramp taxi time prediction accuracy of the flights having those features.

Extending the existing variables used in the pushback time prediction, runway configuration, taxi distance, and the number of departures and arrivals in the ramp area are added to the feature set for the ramp taxi time prediction. CLT has three runway configurations: North, South_Converging, and South_Simultaneous. Since the runway configuration defines spot assignments of departures, as well as traffic flow patterns in the ramp area, it can affect the taxi times inside the ramp. In fact, the average ramp taxi time in North flow (4.02min) is much shorter than in South flow (6.91min). In a low traffic condition, ramp taxi distance from gate to spot is considered as a dominant factor that determines the unimpeded taxi-out time because the travel time is proportional to its distance when a flight moves at a constant speed. The numbers of departures and arrivals taxiing in the ramp area when a departure leaves its gate are also included in the variable list to account for the congestion level on the ground and the taxi delay added to the unimpeded taxi time. Figure 10 illustrates the relationship between the actual ramp transit times and the number of departures and arrivals moving on the ground from the one-month historical data in August 2018. Each linear trendline inside the scatter plots shows a weak positive correlation between ramp transit time and ramp congestion.

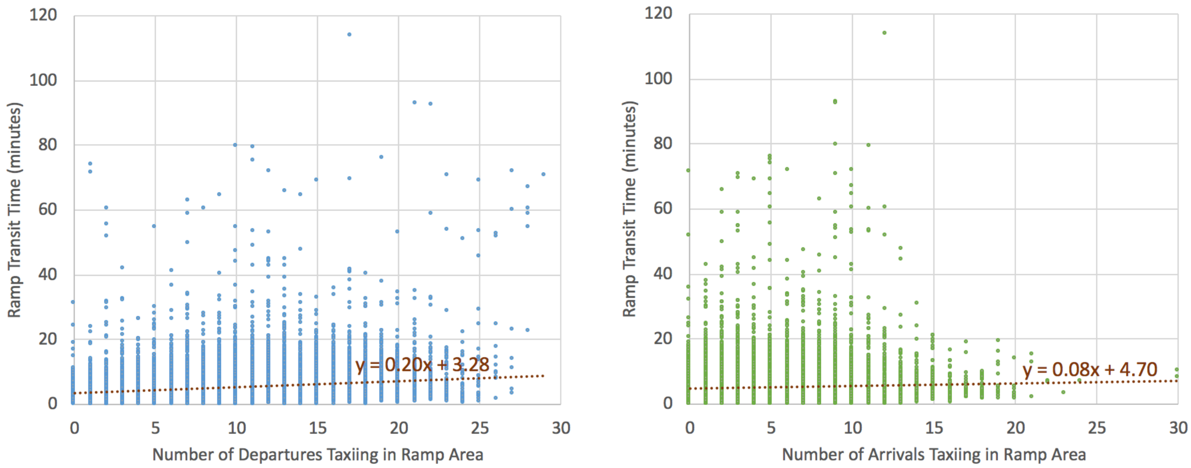

Figure 10. Ramp Transit Times as a Function of Number of Departures (left) and Arrivals (right) Taxiing in Ramp Area

To sum up, a total of 99 features are defined for the ramp taxi time prediction using machine learning algorithms, as described above. These features include 18 binary variables for ramp area (3 for A concourse, 4 for B concourse, 3 for C concourse, 3 for D concourse, and 5 for E concourse), 25 binary variables for spots, 23 binary variables for air carriers, 23 binary variables for aircraft types, 3 binary variables for runway configurations, one binary variable for a potential gate conflict with an arrival, one variable for the pushback time of a day (in hour), and 2 binary variables for TMI (one for APREQ and the other for EDCT). In addition to them, three more features are included for the ramp taxi time prediction, which represent the ramp taxi distance from gate to spot along with the standard taxi route, and the numbers of departures and arrivals taxiing in the ramp area when a departure leaves it gate.

The ramp transit time prediction results from various machine learning algorithms for the same training and test dataset in August 2018 are presented in Table 5. The prediction performance metrics of the data-driven prediction model using the default ramp taxi speed value (6.6 knots), named 'Default' hereinafter, are also added in the last column for direct comparison. As the machine learning algorithms show different rankings depending on what kind of performance metrics is measured, it is hard to tell which algorithm is the best. In general, they show similar prediction performance to each other, but the SVR model shows relatively worse performances in terms of the average of errors, MAE, RMSE, and flights within ±5min error. The Default prediction model also shows a slightly lower prediction accuracy, compared to the machine learning algorithms, but the difference is not significant. This comparison result supports that the main contributor to the ramp taxi time prediction is the taxi distance from gate to spot, and that the benefits from considering other factors such as gate conflict, TMIs, and ramp congestion level may be limited in terms of the prediction performance improvement.

Table 5 Ramp Transit Time Prediction Accuracy Comparison with Various Machine Learning Models

Prediction Model	LR	SVR	Lasso	*k*NN	RF	NN	Default
Average of errors (min)	0.02	1.29	0.00	0.13	-0.01	0.04	0.07
Median of errors (min)	-0.53	-0.01	-0.59	-0.33	-0.49	-0.50	-0.45
Mean Absolute Error (min)	2.08	2.38	2.09	2.17	2.00	2.01	2.32
Root Mean Square Error (min)	3.56	4.36	3.60	3.80	3.54	3.52	4.00
Flights within ±1min error	37.7%	40.2%	37.3%	41.7%	43.0%	41.1%	37.4%
Flights within ±3min error	80.9%	79.2%	80.8%	78.5%	81.7%	81.1%	76.4%
Flights within ±5min error	93.0%	88.4%	93.1%	90.7%	92.5%	93.0%	90.4%

Figure 11 shows the box plot of the prediction error distribution from each prediction method. Because of the heavy tail on the right-hand side of the ramp transit time distribution (see Figure 3), the average value of the prediction error (X mark in each box) is always greater than the median value (the middle line of the box) for all the prediction models. From this plot, it can be seen that the Random Forest (RF) shows the best prediction accuracy, whereas the Support Vector Regression (SVR) mostly predicts shorter taxi times than actual ones. It is also noted that the prediction performance of the Default prediction model is as good as the Linear Regression (LR) model's.

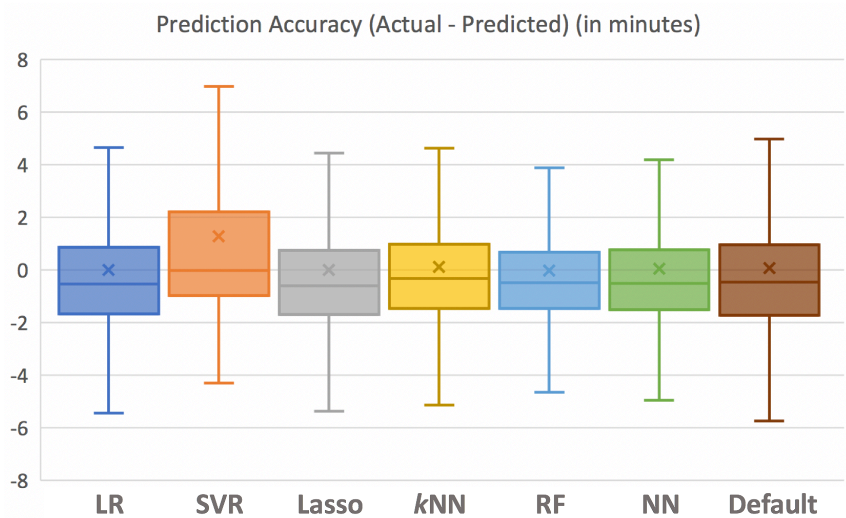

Figure 11. Box plot for ramp transit time prediction accuracy from various machine learning algorithms

V. Conclusions

In this paper, we analyzed the ramp taxi time data at CLT and developed the data-driven prediction models for the pushback process time and the ramp transit time. Leveraging accurate surveillance and user input data from the ATD-2 systems deployed at CLT, we are able to divide the ramp taxi time into the pushback process time from pushback until taxi clearance approval and the pure ramp transit time from gate to spot. By investigating the variables impacting the pushback time, a simple prediction model based on the decision tree having two main criteria, ramp area and aircraft type, was developed for fast pushback time prediction. For the ramp transit time prediction, a separate prediction model using the default ramp taxi speed value and the taxi distance from gate to spot was introduced for the use in real-time operations at the field. For the evaluation, several machine learning algorithms were applied to the same dataset from the actual operations data at CLT. Even though the machine learning algorithms tried more variables that could affect the pushback and ramp transit times, the data-driven prediction models showed similar prediction performances for both pushback time and ramp taxi time predictions.

Considering the large amount of uncertainties in flight operations on the airport surface, it is not possible to make a perfect prediction for the pushback time and taxi time of individual departure flights. Nevertheless, if similar flight data quality becomes available at other airports, this data-driven approach could be applied to the pushback and taxi time predictions there. Depending on the operational characteristics and the geometry of an airport layout, different decision criteria may be applied for the target airport to achieve better prediction accuracy. Statistical approaches like principal component analysis (PCA) would be helpful for determining those decision criteria.

References

[1]Aponso, B., Coppenbarger, R., Jung, Y., O'Connor, N., Lohr, G., Quon, L., and Engelland, S., "Identifying Key Issues and Potential Solutions for Integrated Arrival, Departure, Surface Operations by Surveying Stakeholder Preferences," AIAA 2015-2590, *the 15th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Dallas, TX, June 22-26, 2015.

[2]Jung, Y., Engelland, S., Capps, A., Coppenbarger, R., Hooey, B., Sharma, S., Stevens, L., Verma, S., Lohr, G., Chevalley, E., Dulchinos, V., Malik, W., and Morgan Ruszkowski, L., "Airspace Technology Demonstration 2 (ATD-2) Phase 1 Concept of Use (ConUse)," NASA/TM-2018-219770, February 28, 2018.

[3]Ging, A., Engelland, S., Capps, A., Eshow, M., Jung, Y., Sharma, S., Talebi, E., Downs, M., Freedman, C., Ngo, T., Sielski, H., Wang, E., Burke, J., Gorman, S., Phipps, B., and Morgan Ruszkowski, L., "Airspace Technology Demonstration 2 (ATD-2) Technology Description Document (TDD)," NASA/TM-2018-219767, March 31, 2018.

[4]Sharma, S., Capps, A., Engelland, S., Jung, Y. C., and Coppenbarger, R. A., "Operational Impact of the Baseline Integrated Arrival, Departure and Surface Scheduling Field Demonstration," *the 37th Digital Avionics Systems Conference (DASC)*, London, England, UK, September 23-27, 2018.

[5]Coupe, W. J., Bagasol, L. N., Chen, L. X., Lee, H., and Jung, Y. C., "A Data Driven Analysis of a Tactical Surface Scheduler," *2018 Aviation Technology, Integration, and Operations (ATIO) Conference*, Atlanta, GA, June 25-29, 2018.

[6]Idris, H. R., Clarke, J. P., Bhuva, R., and Kang, L., "Queuing Model for Taxi-Out Time Estimation," *Air Traffic Control Quarterly*, Vol. 10, No. 1, 2002, pp. 1-22.

[7]Simaiakis, I., and Balakrishnan, H., "Queuing Models of Airport Departure Processes for Emissions Reduction," *AIAA Guidance, Navigation, and Control (GNC) Conference*, Chicago, IL, August 2009.
[8]Badrinath, S., Balakrishnan, H., Clemons, E., and Reynolds, T. G., "Evaluating the Impact of Uncertainty on Airport Surface Operations," *2018 Aviation Technology, Integration, and Operations (ATIO) Conference*, Atlanta, GA, June 25-29, 2018.
[9]Legge, J., and Levy, B., "Departure Taxi Time Predictions Using ASDE-X Surveillance Data," *the 26th Congress of International Council of the Aeronautical Sciences (ICAS)*, Anchorage, AK, September 14-19, 2008.
[10]Kistler, M. S., and Gupta, G., "Relationship between Airport Efficiency and Surface Traffic," *the 9th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Hilton Head, SC, 2009.
[11]Jordan, R., Ishutkina, M., and Reynolds, T., "A Statistical Learning Approach to the Modeling of Aircraft Taxi Time," *in Proceedings of the IEEE/AIAA 29th Digital Avionics Systems Conference (DASC)*, Salt Lake City, UT, October 3-7, 2010.
[12]Srivastava, A., "Improving Departure Taxi Time Predictions Using ASDE-X Surveillance Data," *the 30th Digital Avionics Systems Conference (DASC)*, Seattle, WA, October 16-20, 2011.
[13]Ravizza, S., Atkin, J. A. D., Maathuis, M. H., and Burke, E. K., "A Combined Statistical Approach and Ground Movement Model for Improving Taxi Time Estimations at Airports," *Journal of the Operational Research Society*, Vol. 64, 2013, pp. 1347-1360.
[14]Ravizza, S., Chen, J., Atkin, J. A. D., Stewart, P., and Burke, E. K., "Aircraft Taxi Time Prediction: Comparisons and Insights," *Applied Soft Computing*, Vol. 14, 2014, pp. 397-406.
[15]Chatterji, G. B., and Zheng, Y., "Wheels-Off Time Prediction Using Surface Traffic Metrics," *the 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Indianapolis, IN, September 17-19, 2012.
[16]Lee, H., Malik, W., Zhang, B., Nagarajan, B., and Jung, Y. C., "Taxi Time Prediction at Charlotte Airport Using Fast-Time Simulation and Machine Learning Techniques," *the 15th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Dallas, TX, June 22-26, 2015.
[17]Lee, H., Malik, W., and Jung, Y. C., "Taxi-Out Time Prediction for Departures at Charlotte Airport Using Machine Learning Techniques," *the 16th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Washington D.C., June 13-17, 2016.
[18]Coupe, W. J., Milutinovic, D., Malik, W., and Jung, Y. C., "A Data Driven Approach for Characterization of Ramp Area Push Back and Ramp-Taxi Processes," *the 16th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Washington D.C., June 13-17, 2016.
[19]Scikit-learn Library for Machine Learning in Python, Version 0.20.0, 2018. http://scikit-learn.org/stable/tutorial/index.html