

Improved Prediction of Gate Departure Times Using Pre-Departure Events

Lara S. Cook* and Stephen Atkins†
Mosaic ATM, Inc., Leesburg, VA, 20175

and

Yoon Jung‡
NASA Ames Research Center, Moffett Field, CA, 94035

Traffic management decisions and automation supporting those decisions currently lack accurate departure demand information. Future departure demand is typically predicted using the scheduled departure times of each individual flight, which are shown to be poor estimates of actual departure times. This paper describes an approach to improve the prediction of parking gate departure times for individual flights, thereby improving the knowledge of departure demand at longer prediction horizons. The approach uses air carrier-provided data about when certain milestones in preparing for departure are completed. Actual times for pre-departure events are compiled in advance from historical data. These statistics are then used to improve gate departure time predictions in real-time. This paper presents an algorithmic approach for applying pre-departure event times to improve gate departure time predictions for individual flights and applies the method using Aircraft Communication Addressing and Reporting System (ACARS) messages from a large domestic air carrier at a busy airport. Results show the benefit of applying this pre-departure information to two existing automation systems that predict gate departure times – the Surface Management System and the Enhanced Traffic Management System. The application of ACARS data using the proposed algorithm provides a 36% improvement to the gate departure time predictions.

I. Introduction

TRAFFIC management decisions and automation supporting those decisions currently lack accurate departure demand information. Future departure demand is typically predicted using the scheduled departure times of individual flights. These times have been shown to be poor estimates of actual departure times, especially during periods of high delays when traffic management is most necessary.¹ Errors in predicted departure demand affect national flow management as well as traffic management at the regional – Air Route Traffic Control Center (ARTCC) and Terminal Area Approach Control (TRACON) – and airport level.

A flight's actual takeoff time is revealed when the aircraft takes off and is detected by terminal area radar. Prior to discovering that the flight is airborne, departure demand is predicted first using scheduled times from the Official Airline Guide (OAG) and then filed times after the flight operator files flight plan information. This information can be supplemented by airline-provided updates to flight times via the Collaborative Decision Making (CDM) process. Departure predictions are currently inaccurate, as shown later in this paper, with flights departing both early and late relative to the predictions. Improved predictions of when each flight will depart (or will be ready and want to depart) would directly improve traffic management decisions.

For departures from most airports, there is no advance notice when flights will take off earlier than scheduled. When flights take off later than scheduled, only traffic managers in the FAA Air Traffic Control Tower (ATCT) know whether or not the flight has left the parking gate and, therefore, whether the late flight will take off shortly or not for at least the flight's taxi time when it has not yet left the gate. In a few airports, electronic flight strip (EFS)

* Principal Analyst, 801 Sycolin Road, S.E., Suite 212, Leesburg, VA 20175.

† Principal Analyst and Vice President, 801 Sycolin Road, S.E., Suite 212, Leesburg, VA 20175, Senior Member.

‡ Aerospace Engineer, Mail Stop 210-6, Moffett Field, CA 94035, AIAA Member

systems, including the Departure Spacing Program (DSP), provide information to TRACON and ARTCC controllers and traffic managers about which aircraft are moving on the airport surface and their progress toward their departure runways. Research is also in progress under the CDM Surface Management Sub-Team[§] to use airport surface surveillance, such as the Airport Surface Detection Equipment – Model X (ASDE-X), to improve predictions of takeoff times using information about whether or not the flight has left the parking gate and its taxi progress toward a runway.

These techniques provide a somewhat earlier observation of aircraft movement than waiting for the terminal area radar to detect the flight after takeoff. How far in advance the takeoff time prediction is improved depends on the flight's taxi time; the longer the taxi time, the earlier the takeoff time prediction is improved. However, these techniques provide little benefit prior to aircraft leaving their parking gates. Moreover, the required EFS or ASDE-X systems are only available at a small number of airports.

Furthermore, these techniques provide little benefit for airport surface traffic management. On the airport surface, knowledge of departure demand means knowing when aircraft will push back from their parking gates. Whether a flight has left its gate may not be known to controllers and traffic managers in the ATCT until the flight reaches a handoff spot between ramp control and ATCT control. Surface geometry and local agreements about how responsibility is divided between the ATCT and air carrier or airport authority ramp towers or stations can further complicate this prediction. In these cases, surface surveillance systems such as ASDE-X may not provide coverage of the ramps not controlled by the ATCT. At other airports, knowledge that a flight represents actual demand on the airport surface starts when aircraft call for pushback (i.e., clearance to leave their parking gate).

Currently, no advance warning is available when a flight will push back from its gate earlier than scheduled. If a flight is late relative to its schedule, traffic managers in the ATCT usually have no information about how late the flight will be. Air carriers who are CDM participants may modify the estimated gate departure times, but the accuracy of these modifications vary between the air carriers.

II. Approach

This paper describes an approach that improves the prediction of gate push back, or Out times, thereby improving the knowledge of departure demand at longer prediction horizons. Moreover, the approach could be applied at every airport without installing expensive sensors or automation. The approach uses data provided by the air carrier or National Airspace System (NAS) user. These data describe when certain intermediate milestones in preparing for departure are completed. Statistics for when these pre-departure events are expected prior to the actual Out time are compiled in advance from large amounts of historical data. These statistics are then applied in real-time to improve Out time predictions. Results show significant potential benefits. Some uncertainty in a flight's gate departure time is unavoidable, for example if an aircraft experiences an unexpected mechanical problem shortly before planning to push back. In these cases, the predicted Out time error may be very large, though no larger than the error would be under today's systems. The departure predictability for many flights may be improved by using information about the status of the departure preparations.

This paper presents an algorithmic approach for applying pre-departure event times to improve block out, or gate push back, time predictions for individual flights. The paper then describes the application of this method using Aircraft Communication Addressing and Reporting System (ACARS) messages from a large domestic air carrier at a busy airport. ACARS messages that mark several pre-departure events were used in this study. The paper presents results that show the benefit of applying this pre-departure information using the described algorithm to two existing automation systems that predict Out times – the Surface Management System (SMS) and the Enhanced Traffic Management System (ETMS). SMS is a NASA/FAA research airport surface traffic management system.²⁻⁵ ETMS is an FAA automation system widely used for national and regional traffic management.⁶

The algorithm proposed in this paper could be used by an FAA automation system such as the ETMS. Although pre-departure event information exists for many flights, the information is currently not easily available. In addition to demonstrating the value of this information to traffic management systems, and an algorithm for improving departure demand predictions, a goal of this paper is to motivate progress toward a standard method for FAA traffic management systems to gain access to these air carrier data. Currently, a mechanism already exists for flight operators to submit improved gate departure time estimates to ETMS; however, most operators do not have more accurate estimates. Therefore, flight operators could use the proposed algorithm and provide ETMS with improved gate departure time estimates using the existing mechanism.

[§] <http://cdm.fly.faa.gov/Workgroups/surface.html>

Figure 1 illustrates the current process for estimating gate push back time in SMS. Figure 2 captures a possible implementation of the Out time prediction algorithm discussed in this paper. The algorithm could be implemented within SMS, and the improved estimations of gate push back times could be sent to ETMS using the existing mechanism.

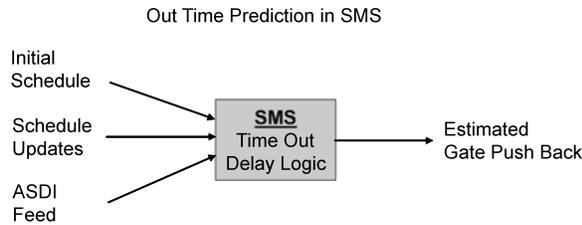


Figure 1. Current process for estimating gate push back times in SMS.

This paper is organized as follows. The next sections describe the data that were used in this study and an analysis of the accuracy of currently available gate departure time predictions. The following sections present the algorithm for applying pre-departure events to improve gate departure time predictions and the process used to select the algorithm parameters. Lastly, results of applying the algorithm to improve SMS and ETMS gate departure time predictions are presented and conclusions and opportunities for future research are listed.

III. Data Collection

The data sources for this analysis were SMS log files from Louisville (SDF) airport and ETMS data. The SMS log files contain both the SMS predictions of gate departure times and the ACARS messages. The ACARS Out message was used for the actual Out time. The SMS log files and the ETMS data both contain the original scheduled data; SMS receives Aircraft Situation Display to Industry (ASDI) data which are a subset of the ETMS data.

A total of 36 SMS log files were used for this study, covering portions of time from June through August, 2006. Typically 4 SMS log files cover one day. ETMS data were processed for the same time periods. This research utilized two tools developed by Mosaic ATM, Inc. The Surface Operations Data Analysis and Adaptation (SODAA) tool ingests SMS log files and raw surface surveillance data into a database and facilitates researchers' construction and visualization of database queries.^{7,8} An ETMS parser and database allows ETMS orig files to be loaded into a database for research queries.

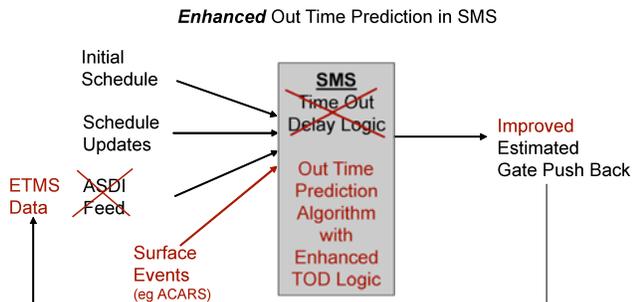


Figure 2. Possible implementation of the out time prediction algorithm within SMS with improved estimated gate push back times being provided to ETMS.

Five pre-departure events – Crew on Board (ACARS Initialization), Crew on Board (Flight Plan Upload), Cargo Door Closed, Load Complete, and Crew Door Closed – were studied. These events were chosen since they are the ACARS events recorded by the particular air carrier prior to gate push back. The times for each of these events were obtained from ACARS messages recorded in the SMS log files. The Crew on Board (ACARS Initialization) message is automatically generated by the aircraft when the pilots power on the ACARS system. This is generally one of the first things the pilots do after entering the aircraft and, therefore, represents a reliable measure of when the crew boards the aircraft. The Crew on Board (Flight Plan Upload) message is generated when the initial flight plan information is received by the aircraft via ACARS. The Door Closed times come from messages automatically sent by the aircraft whenever the aircraft door is closed. The Load Complete message is manually sent and, therefore, susceptible to human error. The Load Complete event is generally expected after the Cargo Door Closed event and before the Crew Door Closed event, but these events may occur in different orders.

Merging the SMS and ETMS data required matching the flights across the two different data sources. For a flight to be used in this analysis, the flight was required to be found in both the SMS and ETMS data. Moreover, the flight must have had at least one Crew on Board, Crew or Cargo Door Closed, or Load Complete ACARS message and have had an ACARS Out message. Lastly, the flight must have had an original scheduled gate departure time in at least one of the data sources. After merging the data sets, a total of 872 flights were studied. The goal was to generate a sufficient flight set, not to perfect the flight matching. A number of flights were discarded that could have been included using different processing. For example, flights split across two SMS log files were discarded in

the present approach. Note that all flights were from a single major air carrier and were departures from the single studied airport.

Several queries were run against the SODAA and ETMS databases to produce tables of data: ACARS messages, SMS Out time predictions, ETMS messages, actual Out times (from the ACARS Out message), and original scheduled Out times. Several simple programs were written to process and merge the data into the format necessary for running the Out time prediction algorithm.

IV. Baseline Data

To measure the improvements that ACARS data can provide, the current prediction accuracy attained without the use of ACARS data is calculated as a baseline. Three current Out time prediction sources are studied –OAG, ETMS, and SMS. The OAG scheduled time is equivalent to the initial ETMS scheduled time and the initial SMS gate time of departure, but the ETMS gate time may subsequently change as a result of CDM messages, flight plans, and time-out delay logic. The following examines these three baselines and compares them to understand which data source provides the best predictions of Out times at various prediction horizons (i.e., amounts of time prior to the actual Out time).

A. OAG Original Scheduled Gate Departure Time

The OAG scheduled departure time, which is a gate departure time, not a scheduled takeoff time, is often used as a prediction of the flight’s block out time. It can be the only prediction of gate departure time for many hours until a flight plan is filed. The first ACARS-provided Out time is used as truth against which the prediction accuracy is measured.**

Figure 3 illustrates the inaccuracy of the scheduled gate push back times for those flights included in this study. Flights rarely push back any earlier than

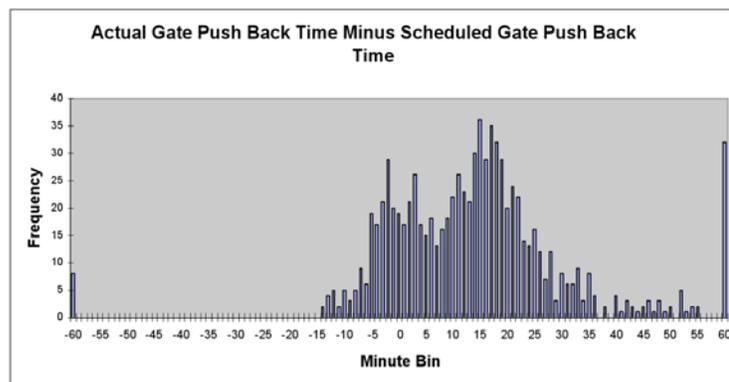


Figure 3. Histogram of prediction errors for scheduled data.

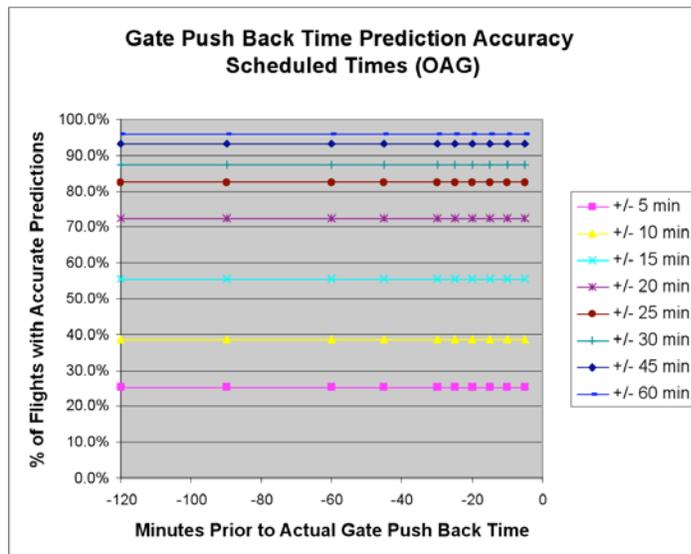


Figure 4. Gate push back time prediction accuracy of scheduled times for various accuracy windows.

15 minutes prior to their scheduled time, but they can push back quite a bit later, as illustrated in the right tail of the distribution. For this data set, flights pushed back on average 14 minutes later than their scheduled times.

Figure 4 plots the percentage of flights for which the OAG scheduled departure time was within a specified accuracy window of the ACARS Out time. The OAG data are static, provided 24 hours or more prior to the scheduled departure time. Therefore, there is no dependence on prediction horizon. However, the graph is drawn as a function of time prior to ACARS Out time for consistency with later graphs. Each data series in the graph represents a prediction accuracy window.

For example, the OAG scheduled departure time is within +/- 5 minutes of the

** A complete analysis of the ACARS data is beyond the scope of this paper. In some situations, more than one ACARS Out message may be sent for a flight.

actual Out time for only 25% of the flights. Wider accuracy windows capture larger percentages of flights in the expected way. But even at the widest accuracy window shown, +/- 60 minutes, not all of the flights are captured. A total of 4.4% of the flights push back more than +/- 60 minutes from their OAG scheduled time.

B. Enhanced Traffic Management System

This section describes the use of the ETMS Estimated Gate Time of Departure (EGTD) as a prediction of Out time. A variety of ETMS messages may be received prior to a flight leaving its parking gate and can modify the EGTD. These messages include:

- F – Filed flight plan
- N – Airline-provided flight modification (CDM participants only)
- B – Time-out delay logic triggered by ETMS once the current time is equal to or later than the estimated runway time of departure
- A – Amended flight plan
- E – Control Times Issued
- Q – Airline-provided flight creation (CDM participants only)

These updates should improve the prediction accuracy at shorter prediction horizons.

Figure 5 plots the percentage of flights for which the ETMS EGTD was within the specified window of the actual Out time, as a function of time prior to actual Out time. The ACARS Out time was again used as the actual Out time. Each data series in the graph represents a prediction accuracy window. For smaller accuracy windows (i.e., +/- 5 minutes and +/- 10 minutes), the prediction accuracy improves as the prediction horizon decreases. However, for larger accuracy windows, the prediction accuracy initially improves with decreasing prediction horizon but then worsens for the shortest prediction horizons.

The underlying flight data were analyzed to determine the cause of the prediction errors worsening in the last 10 minutes prior to the actual Out time. The air carrier used in this study provided new “N” messages, or airline-provided CDM flight modification times, to ETMS close to the actual Out time. These schedule updates tended to predict a gate departure time later than when the flight actually departed. Different carriers likely have different processes and systems for submitting flight modification data to ETMS and, therefore, this characteristic may be isolated to the air carrier studied in this paper. However, our hypothesis based on this observation is that the prediction accuracy using ETMS data is not necessarily a reflection on ETMS processes and logic but rather a reflection on the quality of the data that airlines submit to ETMS.

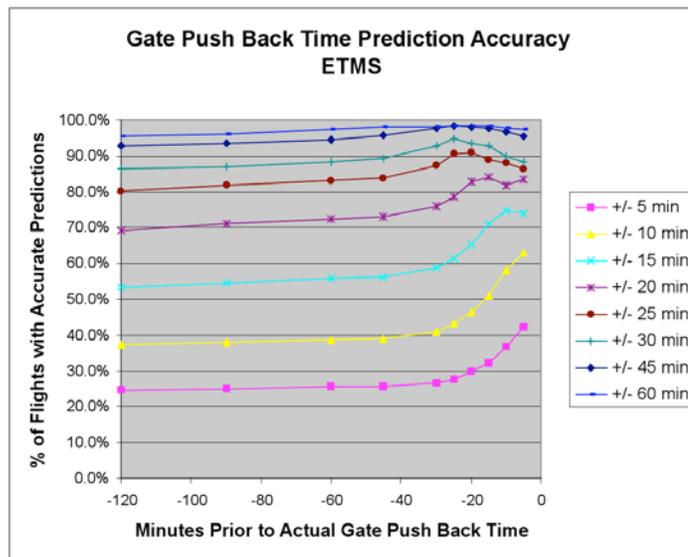


Figure 5. Gate push back time prediction accuracy of ETMS data for various accuracy windows.

Figure 6 plots the percentage of flights for which the SMS estimated Out time was within a specified accuracy window of the ACARS Out time, as a function of the prediction horizon. Out prediction times improve on average as the prediction horizon is reduced. The ASDI data used by SMS do not receive the updates to the ETMS EGTD.

C. Surface Management System

The last data source used as a baseline for Out time predictions was SMS. Like ETMS, SMS relies on a scheduled time for its initial estimate of Out time. SMS receives this scheduled time from ASDI or directly from the air carrier. SMS receives some schedule updates directly from the air carrier. Otherwise, SMS does not modify the estimated Out time unless the current time reaches the estimated Out time and the flight has not yet pushed back. SMS will then adjust the estimated Out time to remain equal to the current time until the flight pushes back. This is typically referred to as time-out delay logic. The SMS time-out delay logic runs every 10 seconds.

Figure 6 plots the percentage of flights for which the SMS estimated Out time was within a specified accuracy window of the ACARS Out time, as a function of the prediction horizon. Out prediction times improve on average as the prediction horizon is reduced. The ASDI data used by SMS do not receive the updates to the ETMS EGTD.

Therefore, the effect of the air carrier’s schedule updates seen in the ETMS performance is not seen in the SMS performance.

D. Comparison of Data Sources

Comparison of the three current data sources for predicting Out times revealed that no one data source provides the most accurate predictions at all prediction horizons. SMS and ETMS both outperform OAG scheduled data. SMS is, on average, always more accurate in predicting Out times than using the OAG scheduled time. SMS improves on OAG scheduled data through the time-out logic and some schedule updates provided by the air carrier to SMS. Since many more flights depart later than scheduled as opposed to earlier, the time-out delay logic is important and makes a significant impact on prediction accuracy at short prediction horizons. The biggest difference between SMS and OAG scheduled data is for the shortest prediction horizons and smallest accuracy windows. SMS predicts almost 60% more flights within the +/-10 minute window than OAG scheduled data alone. However, SMS does have a larger standard deviation of prediction errors and root mean square error (RMSE) than the OAG scheduled data for longer prediction horizons. This is due to poor estimates in the schedule updates provided by the air carrier at longer time horizons.

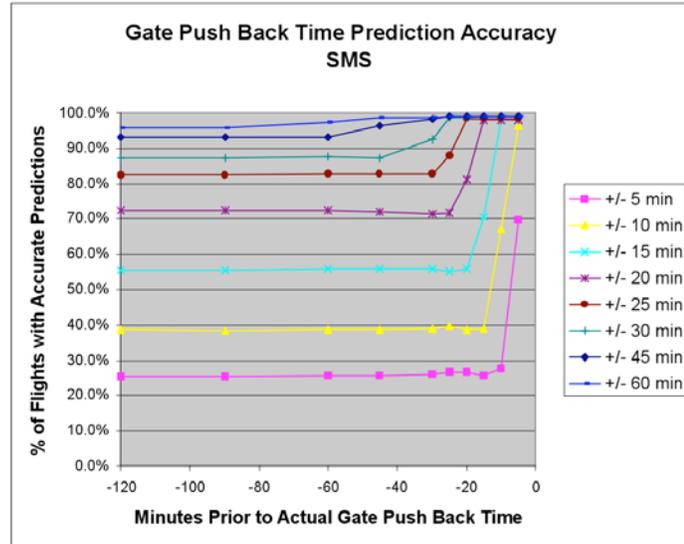


Figure 6. Gate push back time prediction accuracy of SMS data for various accuracy windows.

ETMS consistently provides better predictions (in terms of the percentage of flights within an accuracy window) than the OAG scheduled data for prediction horizons less than one hour. As the prediction horizon increases, the benefit of using ETMS data over just the OAG scheduled data declines. At prediction horizons between 60 and 90 minutes (depending on the accuracy window), the OAG scheduled data become a better predictor than ETMS. Since ETMS starts with the OAG scheduled data, some ETMS data, subsequent to the scheduled data but more than an hour before the actual Out time, must cause this change in prediction accuracy. The ETMS data were studied to find the cause, and it seems to be a result of how the air carrier submits flight modification messages to ETMS.

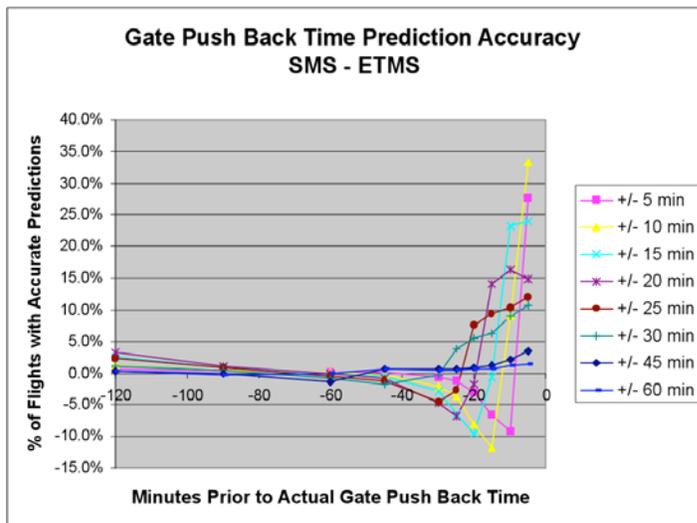


Figure 7. Prediction accuracy differences between SMS and ETMS data.

Though individual CDM ‘N’ messages are sent for flights close to departure time, the air carrier also sends bulk ‘N’ messages for large groups of flights at particular points in time during the day. These are usually for flights that are not expected to depart for many hours in the future. These updates to gate departure times are actually slightly less accurate, on average, than the OAG scheduled gate departure times.

Figure 7 shows the differences in prediction accuracy between SMS and ETMS. A positive value indicates that SMS is more accurate than ETMS for that accuracy window and prediction horizon. Although the results vary with the accuracy window used, the trend is that SMS is, on average, more accurate for prediction horizons greater than 60 minutes and less than 30 minutes. However, for prediction horizons between 30 minutes and 60

minutes, ETMS provides better predictions. The exact prediction horizons at which the crossover occurs depends on the accuracy window.

For longer prediction horizons, SMS outperforms ETMS because SMS is using the original scheduled data while ETMS is using updates that are actually worsening the prediction. For short prediction horizons, SMS outperforms ETMS due to the SMS time-out delay logic. Since many more flights depart later than scheduled as opposed to earlier, the time-out-delay logic is important and makes a significant impact on prediction accuracy at short prediction horizons. Although ETMS also has time-out delay logic, there are two main differences between the SMS and ETMS logic that make SMS more accurate. First, SMS applies the time-out delay logic every 10 seconds; ETMS applies the time-out delay logic every 5 minutes. Consequently, ETMS can predict a flight's Out time as being as much as 5 minutes before the current time. Second, the SMS time-out delay logic is triggered once a flight misses its gate departure time. ETMS does not start applying its time-out delay logic until the flight has missed its runway time of departure. Thus, for SDF, where taxi times are 14 minutes on average^{††}, the SMS time-out delay logic is applied 14 minutes sooner than in ETMS. In the middle prediction horizons, the updated gate times that ETMS receives through messages such as filed flight plans, flight plan amendments, and airline flight modifications – which are not available through the ASDI data that SMS uses – allow ETMS to outperform SMS.

V. Block Out Time Prediction Algorithm

This section describes the Out Time Prediction Algorithm that applies pre-departure event data – ACARS messages in the current study – to improve ETMS and SMS Out time estimates. The algorithm is based on the observation that each type of pre-departure event will occur during a range of time prior to gate departure.

The algorithm is described as follows:

- 1) For each type of pre-departure event, the algorithm assumes the event occurs y minutes prior to actual Out time, where y is a random number described by a normal distribution. The pre-departure events studied in this paper are five ACARS messages: ACARS Initialization Crew on Board, Flight Plan Information Received Crew on Board, Cargo Door Closed, Crew Door Closed, and Load Complete.
- 2) For each event type, a sample set of data is analyzed to estimate the distribution (i.e., mean and standard deviation) for the amount of time prior to the actual Out time that the event is expected to occur.
- 3) A confidence interval is used to define the range of times relative to the actual Out time in which the event is expected to occur. The confidence interval is defined by *min* and *max* values, illustrated in Figure 8. For example, if the mean value is 20 minutes prior to the actual out time, the standard deviation is 5 minutes, and a 95% confidence interval is used, then the minimum value would be 10 minutes (2 times the standard deviation from the mean) and the maximum value would be 30 minutes. These *min* and *max* times are relative to the actual Out time. Note that this approach does not require the distribution be normal; the *min* and *max* values may be selected without requiring symmetry relative to the mean.
- 4) When an ACARS message is received for a flight, the algorithm evaluates whether the current Out time prediction should be changed. The algorithm compares the difference between the event time and the current Out time prediction to the *min* and *max* values for that event type.
 - a) If the current Out time prediction minus the event time is less than (or equal to) *max* and greater than (or equal to) *min*, the current Out time prediction is not changed.
 - b) If the event time is later than *min* before the current Out time prediction, then the Out time prediction is moved later – to be the event time plus *min*, as shown in Figure 9.

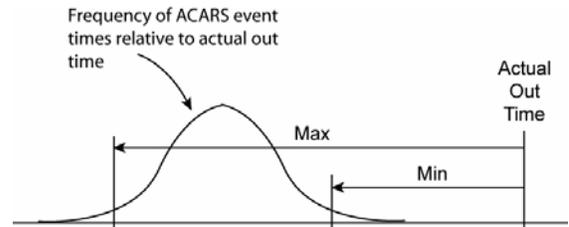


Figure 8. Definition of Min and Max Values.

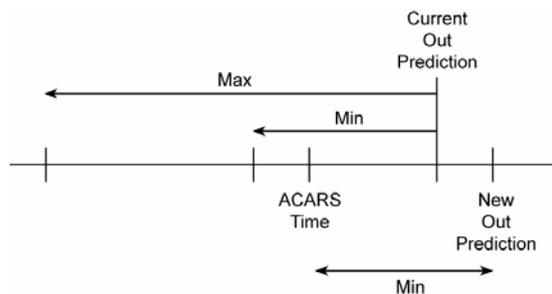


Figure 9. Application of late ACARS message.

^{††} Bureau of Transportation Statistics, <http://www.bts.gov/>

- c) If the event time is earlier than max before the current Out time prediction, then the Out time prediction is moved earlier – to be the event time plus max , as shown in Figure 10.
- 5) The algorithm also uses time-out delay logic, which runs every minute. Once a flight is within x minutes of the currently estimated Out time, the estimated Out time is pushed one minute later. The value of x is dependent on the state of the flight (i.e., what events have already occurred for the flight). For example, flights that are scheduled to operate but for which no ACARS messages associated with pre-departure events have been received will have a larger value for x than flights that have received a Crew on Board ACARS message. The values of x are set based on analysis of the data. Figure 11 illustrates the time-out delay logic.

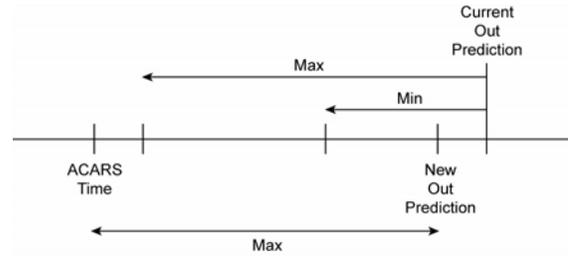


Figure 10. Application of early ACARS message.

For each event type, the algorithm has three parameters – min , max , and x – that must be selected from analyzing historical pre-departure event data. Note that these parameters may be different for different air carriers or airports. Future research will study the robustness of the algorithm to these parameters and the set of characteristics (airport, flight operator, aircraft type, time of day, etc.) that should be used to define sets of parameters. The algorithm was prototyped in Java and tested against the data used to generate the baseline prediction accuracy. The following section describes the process of selecting the algorithm parameters. The subsequent section compares the prediction accuracy using the algorithm with the baseline cases.

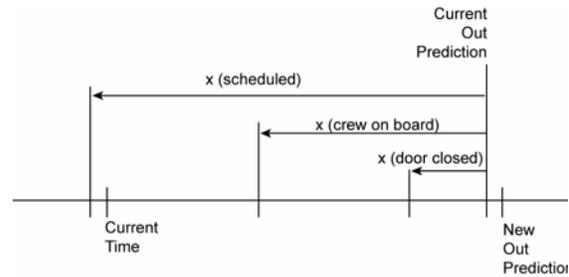


Figure 11. Time-out delay logic.

VI. Identifying Algorithm Parameters

This section discusses the approach used to select the algorithm parameters. Each of the five ACARS message types evaluated as pre-departure events require three parameters: the minimum and maximum bounds on the expected range and the time-out delay parameter. In addition, a time-out delay parameter is needed for use prior to any pre-departure event occurring.

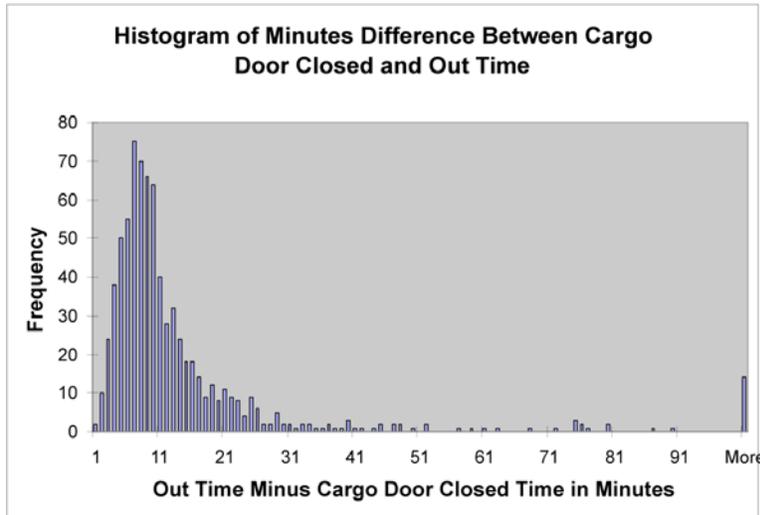


Figure 12. Distribution of actual Out time minus event time for cargo door closed events.

The distribution of the event time relative to the actual Out time was studied for each message type, to understand whether the time of the event relative to the actual Out time is sufficiently consistent and far enough before the actual Out time for the event to be a useful predictor of actual Out time. For example, on average the Cargo Door Closed event occurred 13.5 minutes prior to the actual Out time, with a standard deviation of 17.3. Figure 12 shows a histogram of the differences between the Cargo Door Closed times and actual Out times.

Significant opportunity exists for future enhancements to the algorithm to better handle outliers. In the current analysis, Cargo and Crew Door Closed and Load Complete messages for a

flight were only used if a Crew on Board message had previously been received, which resulted in a “filtered” data set.

Based on analysis of the ACARS data, four of the ACARS message types were selected for use in the algorithm. The ACARS Initialization Crew on Board (M40) message was discarded as having too large a standard deviation. Figure 13 graphs the mean and standard deviation of each of those message types. As expected, the closer an event typically happens to the actual Out time, the smaller the standard deviation.

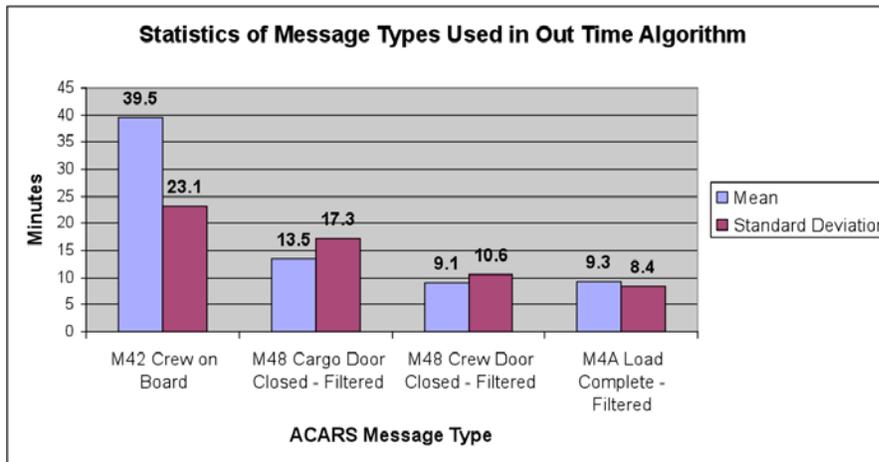


Figure 13. Mean and standard deviation of Out time minus event time by message type.

Several different metrics are useful for evaluating prediction accuracy – the percentage of flights with predictions within n minutes of the actual Out time, the mean of the prediction errors, the median of the prediction errors, the standard deviation of the prediction errors, and the RMSE. The RMSE is a frequently-used measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing

being modeled or estimated.

All of these statistics can be measured at times prior to the actual Out event. Comparing two runs of the algorithm and judging which produced better predictions requires a single scalar statistic. Instead of choosing one prediction horizon, an additional statistic was defined that combined the RMSE at various prediction horizons ($nPH = \text{number of prediction horizons}$), weighting each prediction horizon depending on the relative importance of accuracy in predictions that amount of time in advance. This composite RMSE is defined in Eq. (1).

$$\text{CompositeRMSE} = \sum_{n=1}^{nPH} (\text{RMSE}_n * \text{Weight}_n) \quad (1)$$

The weighting values which were used are shown in Figure 14. The highest weight was given to the time horizon of 10 minutes prior to actual Out as opposed to 5 minutes prior. A good estimate at 10 minutes tends to result in a good estimate at 5 minutes, and our goal was to achieve strong estimates far enough in advance to support improved airport surface planning. Further work includes analyzing the results achieved using different weighting factors.

Minutes Prior to Actual Out	Weight
5	0.25
10	0.5
15	0.1
20	0.05
25	0.04
30	0.03
45	0.02
60	0.01

Figure 14. Weighting values used for the Composite RMSE.

The approach to find the best value for each of the time-out delay parameters was to set all other parameters used by the algorithm to values that would not change the predicted Out time. The other time-out delay parameters were set to 0, as were the *min* values for each event type. The *max* values were set to infinity. The one exception is the *min* parameter for the event type being evaluated for the time-out delay parameter. The *min* parameter should always be greater than or equal to the time-out delay parameter. For the event type being evaluated, the time-out delay parameter was set to various values, running the algorithm for each value. If the event type also had a *min* parameter, the *min* parameter was set to the same value as the time-out delay parameter. The value of the parameter that minimized the composite RMSE statistic was selected.

For each event type, the approach used to find the best values for the *min* and *max* parameters was to set all of the parameters for the other event types to values that would not change the predicted Out time. The time-out delay parameters were set to 0, as were the *min* values for each other event type. The *max* values for each other event type were set to infinity – a value much larger than any possible value of the difference between the actual Out time and an ACARS event time. For the event type being evaluated, the time-out delay parameter was set to the best value found in the prior step. The *min* value was initially set equal to the time-out delay parameter. The value of the *max* parameter was varied, running the algorithm for each value. The value of the *max* parameter that minimized the composite RMSE statistic was selected. Once the best value for the *max* parameter was found, this was held constant and the *min* parameter was varied until the best value was found.

Other methods for selecting the parameters were also explored, either by evaluating the parameters in a different order or by setting previously evaluated parameters to their optimal value instead of 0. Though different methods yield different parameters, when the algorithm was run using these different parameters, the composite RMSEs which resulted were not significantly different (e.g., 17.30 vs. 17.28). This is an area that requires further research in order to determine the optimal method for finding all algorithm parameters.

VII. Results

After determining the best values for each of the algorithm parameters, the Out time prediction algorithm was run using these parameters and the results compared to the baseline cases. Figure 15 graphs the RMSEs (for various prediction horizons) for the original scheduled data alone, the SMS predictions alone, and the SMS predictions improved with ACARS data and the Out time prediction algorithm.

The application of ACARS data through the Out time prediction algorithm improves the Out time predictions at all prediction horizons.

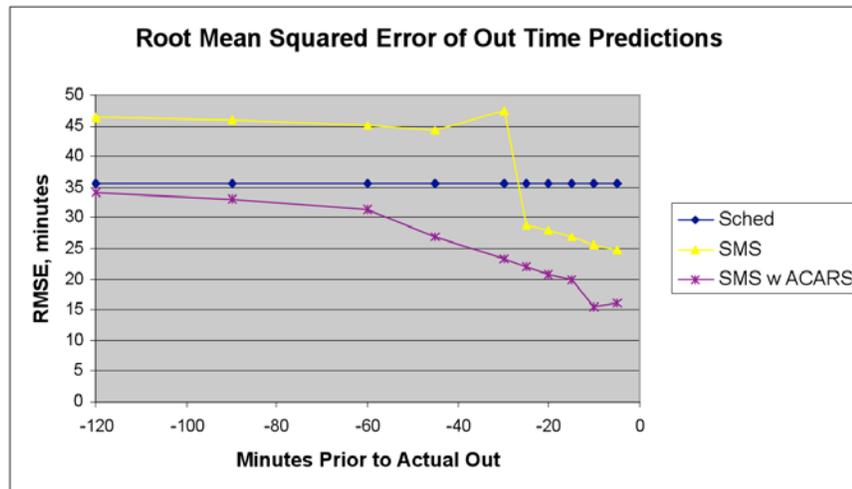


Figure 15. RMSE for SMS vs. SMS with ACARS Data.

Data Source	Composite RMSE
OAG Schedule	64.3
SMS	27.0
SMS w/ ACARS	17.3

Figure 17. Composite RMSE values by data source.

How each individual parameter affects the results of the algorithm has not been studied yet. Future work will include a sensitivity analysis to understand the robustness of the algorithm to parameter choices.

Figure 16 graphs several statistics for each data source – the mean, median, RMSE, and standard deviation of the prediction errors made at 10 minutes prior to the actual Out times. All statistics illustrate an improvement in prediction errors.

Figure 18 graphs the percentage of flights at each

Figure 17 shows the composite RMSE values for each data source. The application of ACARS data using the proposed algorithm provides a 36% improvement to the SMS predictions and a 73% improvement to the OAG Schedule.

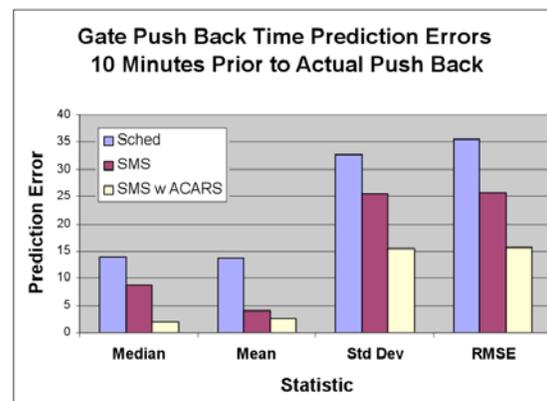


Figure 16. Prediction error statistics at 10 minutes prior to Out.

prediction horizon for which the predicted Out time is within +/-10 minutes of the actual Out time. OAG data provides predictions of Out time within +/-10 minutes for only 40% of the flights for all time intervals. SMS outperforms the OAG data only at the 5 and 10 minute prediction horizons. SMS augmented with the ACARS data provides significant improvement in prediction accuracy for prediction horizons 20 minutes and less and some improvement for prediction horizons between 25 and 35 minutes.

The benefit of applying ACARS pre-departure data to improve ETMS gate departure time predictions was also studied. In addition to the 3 previously discussed baseline cases (the OAG scheduled data alone, ETMS data alone, and the SMS predictions alone), an additional baseline case – SMS and ETMS combined – was created in order to understand the benefits of replacing the ASDI data feed with the ETMS data feed in SMS. In addition to the prior test case (SMS improved with ACARS data), two additional test cases were studied – ETMS improved with ACARS data, and SMS and ETMS combined and improved with ACARS data. The same algorithm parameters that were selected to apply the ACARS data to the SMS data were used for both of the additional test cases. Since those parameters were tuned based on the SMS data, the prediction accuracy results achieved using these parameters with the ETMS data will not necessarily result in the best possible prediction accuracy.

Figure 19 graphs the Out time prediction RMSEs (as a function of prediction horizon) for each of the data sets – the 4 baseline cases and 3 test cases. The SMS with ACARS dataset performs as well or better than all other datasets at every prediction horizon. The addition of ETMS data to the SMS data did not provide any additional benefits.

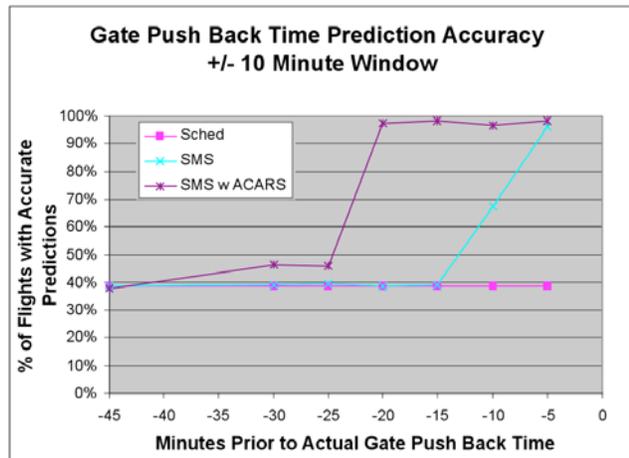


Figure 18. Prediction accuracy within +/-10 minutes window.

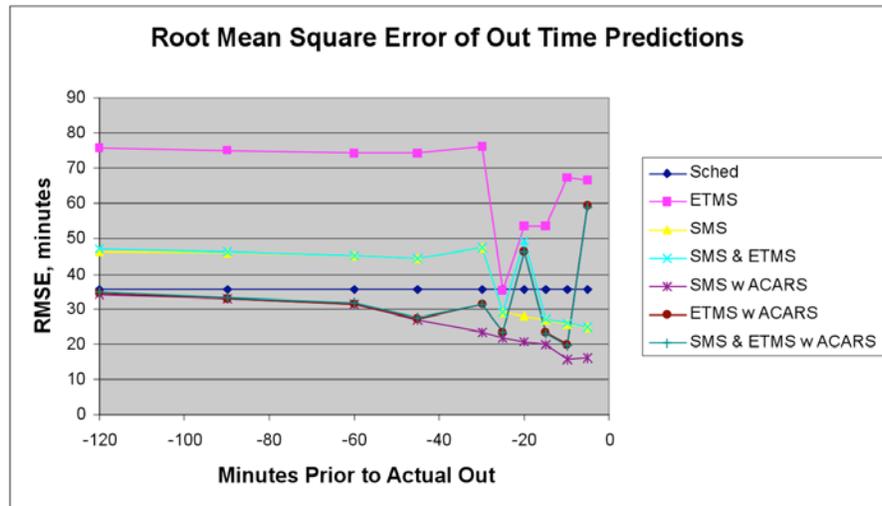


Figure 19. RMSE of Out time predictions over all datasets.

Figure 20 charts the composite RMSEs for each of the datasets. Theoretically, the ETMS data should provide additional benefits to both the SMS and SMS with ACARS data cases. As discussed earlier, the cause was determined to be a characteristic which may be unique to the air carrier from which the present dataset was obtained. The CDM messages sent by the air carrier into ETMS were less accurate than the ETMS data prior to the CDM messages. Different results would be expected if the estimates of Out time provided by the air carrier via CDM messages were more accurate.

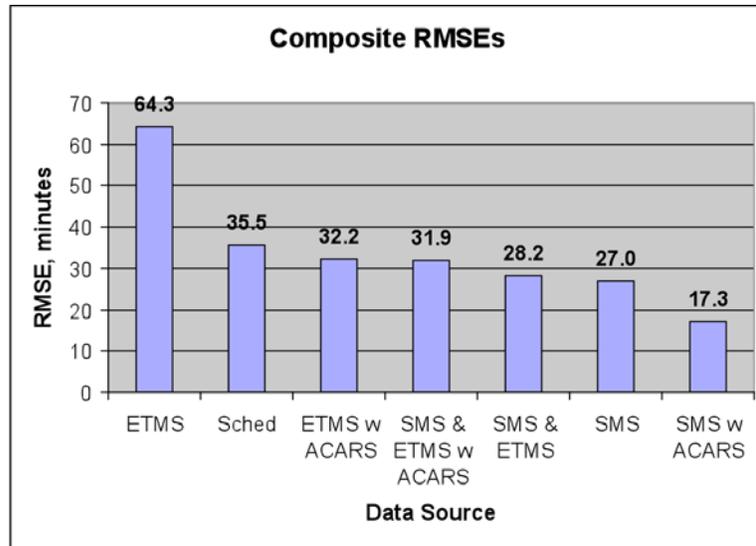


Figure 20. Composite RMSE for each data source.

Figure 21 graphs the percentage of flights, at each prediction horizon for which the predicted Out time is within +/-5 minutes of the actual Out time. The three data sets that apply ACARS data all show a dramatic improvement in prediction accuracy over the data sets that do not incorporate pre-departure event data.

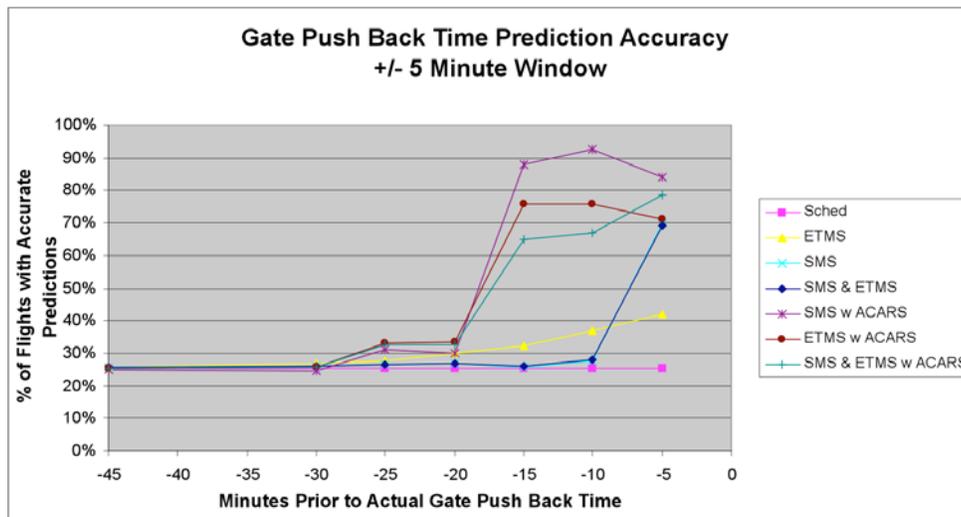


Figure 21. Prediction accuracy within +/-5 minutes window.

VIII. Conclusion

This paper presented an algorithm for using pre-departure event data to improve gate departure time predictions. The methodology using several ACARS messages was applied to improve SMS and ETMS Out time predictions, improving SMS predictions in terms of composite RMSE by 36%. The approach is flexible and could be used with other pre-departure event data and to improve other existing gate departure time predictions. Data from a single flight operator at a single airport was studied in detail. The initial results presented in this paper demonstrate significant potential for the proposed technique.

This research is being continued and a variety of future objectives have been identified. First, we will perform a sensitivity analysis to understand how dependent the prediction improvement is on the algorithm parameters and, moreover, how dependent the optimal parameters are on the data set. As part of this robustness analysis we will study additional data sets for several air carriers at multiple airports. One challenge in wide application of the approach is that although the pre-departure event data exists broadly, it exists in different forms and is not available

from a single source. Future work will include identifying how standard pre-departure events/data is and determining whether other events and possibly other data sources may be more appropriate for some flight operators or airports. For example, pre-departure clearance delivery time obtained from an EFS system may be a useful alternative or additional data source. The method is also not limited to gate departure time predictions. Future work may apply the methodology to improve the prediction of takeoff times at airports that do not have and will not receive detailed surface surveillance systems but that may have EFS.

Acknowledgments

This work was conducted by Mosaic ATM, Inc. under a Phase 2 Small Business Innovation Research (SBIR) contract funded by NASA Ames Research Center.

References

- ¹Shumsky, R.A., "Dynamic Statistical Models for the Prediction of Aircraft Take-off Times," Ph.D. Thesis, Operations Research Center, MIT, Cambridge, MA, 1995
- ²Atkins, S., and Brinton, C., "Concept Description and Development Plan for the Surface Management System," Journal of Air Traffic Control, 2002.
- ³Brinton, C. R., "Collaborative Surface Movement Planning," Final Report of NASA SBIR Contract NAS2-000047, June, 2000.
- ⁴Atkins, S. and Hall, W., "A Case for Integrating the CTAS Traffic Management Advisor and the Surface Management System," AIAA Guidance, Navigation, and Control Conference, Denver, CO, August, 2000.
- ⁵Brinton, C., Krozel, J., Capozzi, B., and S. Atkins, "Improved Taxi Prediction Algorithms for the Surface Management System", Proceedings of the 2002 AIAA Guidance, Navigation, and Control Conference, Monterey, CA, August 2002.
- ⁶Medeiros, M.F., "Advanced Traffic Management System Automation," Proceedings of the IEEE, Vol. 77, Issue 11, Nov. 1989, pp 1643-1652.
- ⁷Brinton, C. R., Cook, L. S., and Atkins, S. C., Collaborative Airport Surface Metering for Efficiency and Environmental Benefits, Proceedings of the Integrated Communication, Navigation and Surveillance Conference, Dulles, Virginia, 2007.
- ⁸Brinton, C. R., Atkins, S. C., and Hall, A., Analysis of Taxi Conformance Monitoring Algorithms and Performance, Proceedings of the Integrated Communication, Navigation and Surveillance Conference, Dulles, Virginia, 2007.