

Impact of Departure Prediction Uncertainty on Tactical Departure Scheduling System Performance

Alan Capps* and Kevin Day*
Mosaic ATM, Fort Worth, TX, 76155

Edward Walenciak†
CSC, Rockville, MD 20850

Shawn A. Engelland‡
NASA Ames Research Center, Fort Worth, TX 76155

Although tactical departure scheduling is commonly used in the National Airspace System, departure prediction uncertainty has a big impact on overall system performance. This paper analyzes the predictive accuracy of individual departure events as observed during operational evaluation of an integrated tactical departure scheduling system at DFW airport. Data from a major air carrier are utilized to improve pushback predictions as well as to notify surface automation of the actual flight pushback time from the gate. Surface event predictive accuracy is analyzed including ramp taxi time predictions, spot crossing duration, airport movement area taxi time predictions, departure clearance and departure runway roll time. Airborne predictive accuracy includes the transit time through the departure fix and to the meter point. Error estimates in the presence of departure uncertainty are used to identify and prioritize areas of future work. Specific improvements implemented during this research are discussed including modifications to airborne departure routing in DFW North Flow Configuration which reduced average error of flight time predictions from 300 seconds to less than 30 seconds. Results of the overall integrated departure system performance are discussed including improved take-off time compliance from an average absolute error 108 seconds to less than 59 seconds.

I. Introduction

In current day traffic management operations, tactical departure scheduling occurs approximately 3.5 times more frequently than departures subject to strategic traffic management initiatives.¹ High tactical departure delays can propagate to national-scale traffic flow challenges that require intervention with strategic procedures like TMA Flow Programs (TFPs).² NASA has developed and demonstrated a tactical departure scheduling system, called Precision Departure Release Capability (PDRC), which provides trajectory-based takeoff time predictions from a Tower automation system for the Center to use when scheduling departures into constrained En Route flows. PDRC also provides for automatic coordination between the Tower and Center to improve upon the efficiency and performance of the current day manual process.³

The PDRC concept and prototype software was developed and evaluated in an operational environment at the Dallas/Fort Worth (DFW) airport at the NASA North Texas Field Site (NTX). The evaluation was conducted from April 30, 2012 to July 26, 2012 and culminated in operational scheduling of 120 flights with the PDRC system. The PDRC system used for this evaluation consisted of the Surface Decision Support System (SDSS) which was utilized at TMC positions in DFW's East and West Towers, and the Research Traffic Management Advisor (rTMA) En Route Departure Capability (EDC) which was utilized by TMC personnel at Fort Worth Center. Additional information on the PDRC concept and prototype system can be found in Ref. 3.

* National Airspace System Engineer, NASA/FAA North Texas Research Station, AIAA Senior Member

† Software Engineer, CSC, AIAA Member

‡ Aerospace Engineer, NASA/FAA North Texas Research Station, AIAA Senior Member

Previous research¹ has shown that tactical departure scheduling is commonly used throughout the NAS; however, little is known about the impact or sensitivity of departure prediction uncertainty on overall tactical departure scheduling system performance. The uncertainties that challenge the success of this integrated system begin on the airport surface prior to pushback, through the flight's transition in the ramp area, to the controlled spot, taxiing to the departure queue, and extending airborne through Terminal Radar Approach Control (TRACON) airspace to crossing of the constrained meter point in en route airspace.

This analysis looks at relevant departure events in tactical scheduling, quantifies the uncertainty associated with each event, and analyzes the achievable level of accuracy in the current day PDRC system. System performance is assessed using a combination of quantitative and qualitative measures obtained during the operational evaluation. Correlations between individual departure event uncertainty and system performance are discussed.

II. Departure Event Uncertainty

To obtain quantitative measures, the relevant departure events, sources of uncertainty and specific measures required for evaluation of tactical departure scheduling were identified. The relevant departure events and their measures are illustrated in Fig. 1.

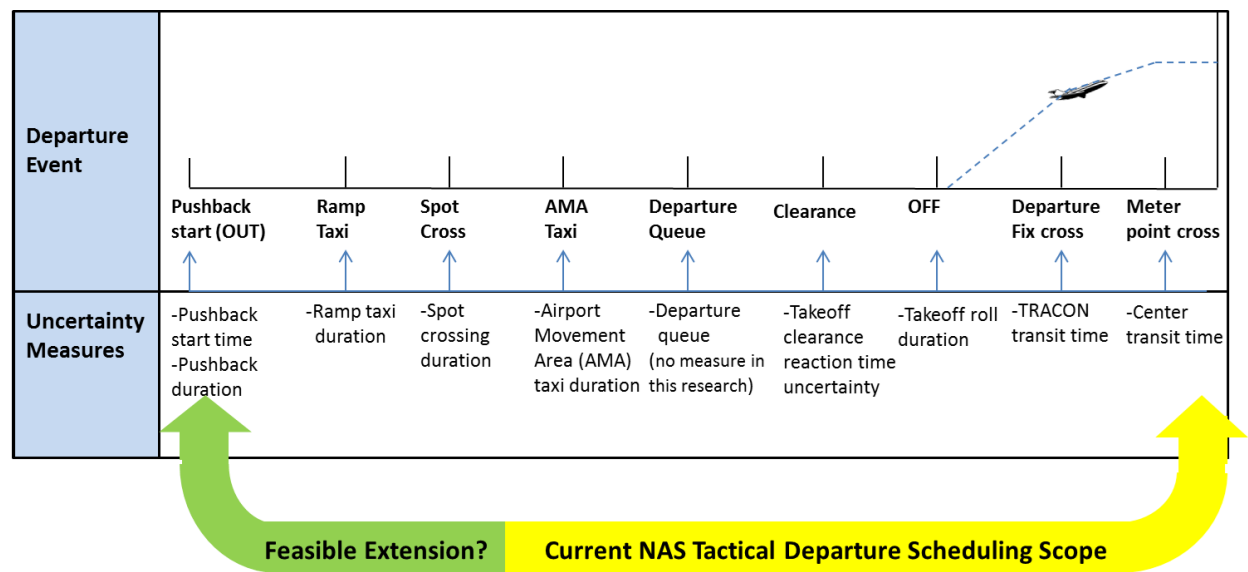


Figure 1. Tactical departure scheduling is affected by the cumulative uncertainty of numerous departure events.

In the National Airspace System (NAS) today, the tactical departure scheduling process¹ generally begins when the flight arrives at the spot (i.e. Apron Entry/Exit Point), and ends when the flight crosses the meter point in Center airspace, which often lies near a neighboring Center boundary. The yellow arrow in Fig. 1 encompasses the departure events that lie within the current-day scope. The green portion of the arrow highlights departure events that precede the current-day scope of tactical departure scheduling. The potential benefits of this extension could include ground delay reduction, airborne delay reduction, lower fuel consumption, and greater situational awareness of the airlines to tactical delay. However, extending the scheduling horizon also adds uncertainty associated with the departure events above the green arrow to the overall predictive accuracy of the estimate.

Data for this analysis were captured during the operational evaluation of PDRC at DFW. A number of specific measures were identified to assess the uncertainty of each element as well as measure the current level of predictive accuracy (or error). Each of these measures is discussed in greater detail in the following sections.

A. Pushback Start Time

The purpose of the pushback start measure was to assess the uncertainty of the predicted pushback time. The SDSS component of the PDRC prototype system begins computing surface trajectories and takeoff time predictions

as soon as a flight plan is received for a given flight. These surface trajectories begin with at the departure gate position and with an estimated pushback (i.e. OUT) time. Multiple sources of pushback time estimates were available to the PDRC system during the evaluation. These include the filed flight plan time, pushback times received directly from an interface with a major air carrier, and a secondary source of pushback data from a commercial flight data service. The primary focus of this research was on the accuracy of pushback received directly from the air carrier.

1. Measurement Approach

The pushback prediction used in this analysis was recorded immediately prior to the actual pushback event. Truth data used to assess the predictive accuracy of the pushback event was the actual out event data received from a gate docking system at DFW airport. This system uses video surveillance of the gate to detect when the flight has moved one meter, at which point it sends notification of the event. The actual OUT data were supplied by the airline system.

Estimated OUT times from the air carrier were compared with actual OUT times from the carrier for all flights from the May 30, 2012 through June 22, 2012. The data used were from air carrier predictions that immediately preceded the actual OUT occurrence, generally within 10 seconds prior to actual OUT event. All flights which did not have an actual OUT or predicted OUT from the airline were removed from the sample. Given the objective was to evaluate the general uncertainty associated with this event and observers were not available to capture details on each OUT event occurrence, the outliers were removed from the sample by removing 1.5 times the interquartile range (IQR). The IQR method of filtering was used due to asymmetric distributions for some departure events in this research. The remaining set of data contained 30,792 departing flights from DFW with air carrier data.

2. Uncertainty Results

The results of the pushback time measure are illustrated in Fig. 2. The mean for this sample was -52 seconds, indicating that flights pushback earlier than their last predicted OUT time from the carrier. After eliminating outliers, the median for this sample was 0 seconds and the standard deviation was 148 seconds. Fifty-three (53) percent of flights in this sample push back earlier or exactly on time with their estimated pushback time. Eighty-six (86) percent of flights push back within one minute late, and ninety-two (92) percent of flights push back prior to 2 minutes late. For tactical departure scheduling purposes, flights that arrive at the spot early pose no challenge for scheduling into the overhead stream. While this is an encouraging statistic, 2,355 flights in this sample pushed back greater than two minutes later than their predicted time. This is significant because if any of these flights were tactically scheduled prior to pushback, it is unlikely the flights could be expedited to depart within the current-day tactical window of two minutes early through one minute late (-2/+1).

A notable observation is that the data received from the airline to report actual OUT time is currently at the resolution of one minute. For error measurement purposes, the actual OUT was compared to the predicted OUT estimate without rounding or truncations of the data. A later version of the airline data feed will contain the actual OUT in seconds level precision.

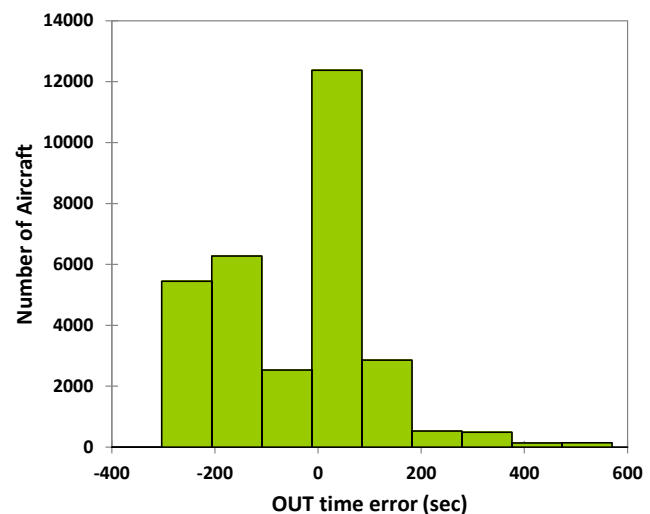


Figure 2. Comparison of airline provided actual OUT with airline provided estimated pushback times.

B. Pushback Duration

The purpose of the pushback duration measure was to assess the uncertainty associated with the range of the pushback event duration. Currently, PDRC's SDSS component accounts for pushback duration with an adaptable "pushback buffer" value. This value may be tailored to account for gate location and ramp geometry. The pushback duration results are utilized in later sections of this paper to determine the pushback duration prediction error. The pushback duration prediction error is the difference in actual pushback duration from the mean pushback duration.

1. Measurement Approach

Often the pushback is thought of as an instantaneous event in time, when in reality it has a duration which is dependent upon a number of factors like jet blast policy for an airport and gate geometry. The pushback event was defined to be the number of seconds from the OUT event discussed in the previous section until the flight begins forward motion under its own power. Observations indicate that nearly all flights at DFW employ the use of a tow tractor, known as a ‘tug’. Truth data used to measure pushback duration are from manual observations made by test personnel.

Manual observations were collected for 194 flights departing various gates during the months of May through July 2012. The observer of the pushback event had either direct line of site visibility to the flight or live video camera with pan/tilt/zoom control capability. The start and end of the pushback event were recorded at second’s level precision. For 138 of the events, the time at which the tug was disconnected was also captured.

2. Uncertainty Results

The results of the pushback duration measure are depicted in Fig. 3. The sample taken had a mean value of 202 seconds, with a median of 189 seconds. The sample demonstrated a significant amount of variation with a standard deviation of 73 seconds, a minimum pushback time of 77 seconds and a maximum pushback time of greater than 7 minutes. The pushback duration data best fit a lognormal distribution. This lognormal distribution is not surprising given this measurement is of time durations which are all positive and independent of one another.

Gate-specific pushback duration variability can occur due to limited space for the tug to push the flight into regions of the ramp taxi area, or areas in which the jet blast from engine start may not be allowed. Due to this geometry, the tug may be required to push the flight back then subsequently pull the flight to a different location prior to disconnecting and aircraft engine start. Significant pushback duration variance existed by air carrier as well with the highest carrier average pushback of 246 seconds and the lowest 148 seconds.

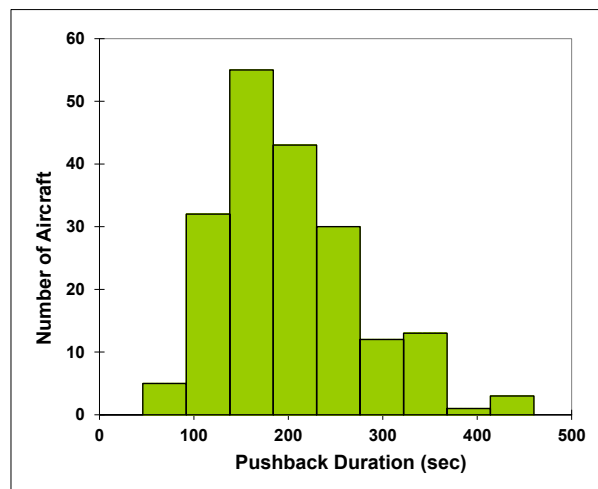


Figure 3. Histogram showing pushback durations.

C. Ramp Taxi Duration

The purpose of the ramp taxi duration measure was to assess the uncertainty associated with the range of ramp taxi time duration. Currently PDRC’s SDSS component models ramp movement as a constant-speed taxi from the gate location to the spot. The ramp taxi duration results are utilized in later sections of this paper to determine the ramp taxi prediction error. The ramp taxi prediction error is the difference in actual ramp taxi time versus the predicted ramp taxi time.

1. Measurement Approach

The ramp taxi event was defined as the number of seconds from forward motion in the ramp area until the aircraft reaches the spot. The primary source of truth data used to assess ramp taxi uncertainty was from manual observations by test personnel with visual access to the pushback event mentioned in the previous section through the flight’s arrival at the spot. Manual observations were collected for 189 flights at various DFW gates during the months of May through July 2012. The start and end of the pushback event were recorded at seconds level precision.

2. Uncertainty Results

The results of the ramp taxi measure are depicted in Fig. 4. The mean ramp taxi time for the sample was 85 seconds with a median of 81 seconds. Some variation was noted with a standard deviation of 40 seconds, and a minimum of 5 seconds with a maximum of over 5 minutes.

The distance from the ramp taxi start location to the spot was an important consideration. The average ramp taxi speed was computed for each flight by dividing the ramp taxi duration by the ramp taxi distance. The mean and median ramp taxi speed for all flights was 8 knots.

A significant amount of ramp taxi speed variation existed within the sample with a standard deviation of 3 kts, a high of 20 kts and a low of 3 kts. The variation was more evident amongst aircraft type than by air carrier, with the lowest average ramp taxi speed of 6.5 kts by the Boeing 737 series and the highest average ramp taxi speed of 8 kts by the McDonnell Douglas MD-80 series.

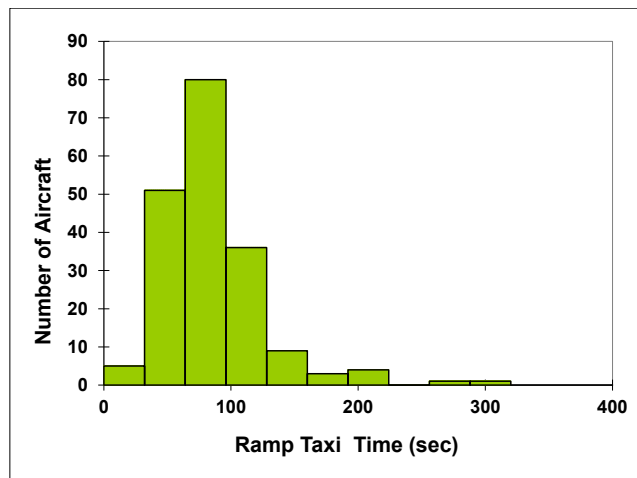


Figure 4. Ramp taxi time duration uncertainty.

D. Spot Crossing Duration

The purpose of the spot crossing duration measure was to measure uncertainty associated with the range of times aircraft were held at the spot prior to entering the airport movement area (AMA). Currently PDRC's SDSS component has the ability to add delay for spot crossing to deconflict with other flights, however, this was not used in the PDRC operational evaluation. Therefore, any time spent waiting at the spot is considered to be prediction error.

1. Measurement Approach

The spot crossing duration was defined as the number of seconds that the flight waited at the airport spot prior to entering the AMA. Manual observations were collected for 190 flights during the months of May through July 2012.

2. Uncertainty Results

The results of the spot crossing duration times are depicted in Fig. 5. Approximately 81% of flights at DFW did not stop at the spot prior to entering the AMA. This can be seen in Fig. 5 by the large number of aircraft which had 0 to 10 seconds delay.

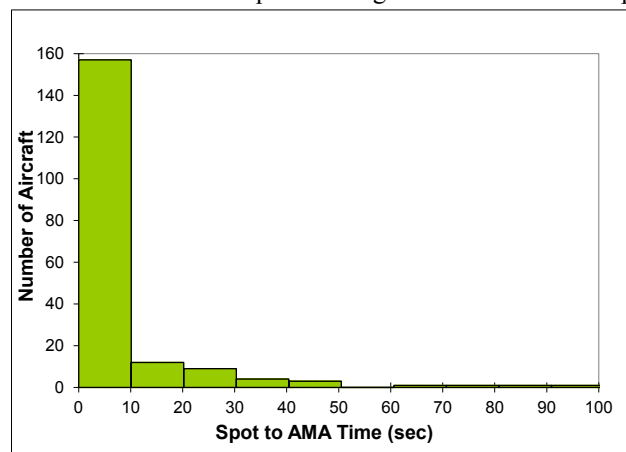


Figure 5. Spot to airport movement area duration.

When flights do stop at the spot, there is generally a small time expense to this action and the average wait is 29 seconds. Significant variation exists with a minimum of 0 seconds and max of 100 seconds for this sample. Many of the longer wait times at the spot can be explained by already present transiting flight on the AMA taxiway. Another situation that may explain the non-zero wait time is ground controller delay in contacting the flight and issuing a clearance to enter the AMA.

E. Airport Movement Area Taxi Duration

The purpose of the AMA taxi duration measure was to assess the uncertainty of the taxi time duration in the FAA-controlled airport movement area. Currently PDRC's SDSS component models AMA movement as a constant-speed taxi from the spot to the runway departure queue following a node-link surface trajectory. AMA taxi prediction error is any error generated by the PDRC system using the current day algorithms and AMA taxi decision trees available for this prediction.

1. Measurement Approach

The AMA taxi time duration is defined as the amount of time from entering the AMA to the point at which the flight enters the departure queue. The AMA taxi time measure intentionally excludes uncertainty associated with the departure queue itself.

Truth data used for AMA taxi time was obtained from post-analysis routines of PDRC output. This logic determined the entry into the AMA as well as the time the flight entered the departure queue. The AMA entry and departure queue time were compared to create the actual ramp taxi time for each flight. The PDRC AMA taxi time prediction was obtained on each flight by assuming a constant AMA taxi speed of 17 knots over the distance between the spot and departure queue which is used in SDSS. For each AMA taxi time prediction, the predicted taxi time to the departure queue was compared with the actual. Outliers were eliminated from the data sample using 1.5 times the IQR. The remaining sample of 46,325 flights from June and July 2012 are discussed in the next section.

2. Uncertainty Results

The results of the airport movement taxi time statistics are depicted in Fig. 6. The overall mean AMA taxi time prediction error is 25 seconds, while the median error is 23 seconds. A positive error indicates the tendency of the PDRC system is to under predict the amount of time it takes for the flight to taxi from the spot to the runway threshold. Under predicting the AMA taxi time is undesirable in tactical departure scheduling because it may allow

insufficient time for the local controller to stage the flight in the departure queue and control the flight to meet its coordinated departure time.

The absolute AMA taxi error from the PDRC system was 35 seconds with a median absolute error of 29 seconds. The average absolute error is approximately 16% of the size of the flight's total AMA taxi time duration.

Variance also existed in the AMA taxi prediction error with a standard deviation of 38 seconds, a minimum of 76 seconds early and a maximum taxi error of 131 seconds late.

The variance is most prominent when viewing the data by air carrier and aircraft type. The top carrier and aircraft type combination have an average of 54 seconds taxi average error, while the lowest have an -13 seconds average taxi error. This suggests that the prediction error could be reduced by using different taxi speeds based upon air carrier and aircraft type.

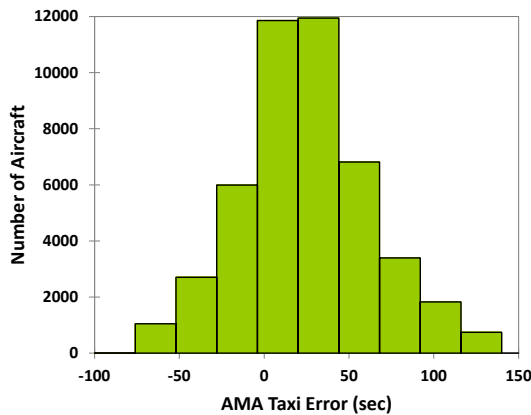


Figure 6. Airport movement area taxi prediction error assuming 17 knot taxi speed.

F. Takeoff Clearance Reaction Time Uncertainty

The purpose of this measure was to assess the uncertainty associated with the range of times of pilot throttle-up response to ATC takeoff clearance. Currently PDRC's SDSS component does not explicitly model this portion of the departure. The takeoff clearance reaction time results are utilized in later sections of this paper to determine the prediction error that would exist assuming the mean clearance reaction time for all flights.

1. Measurement Approach

Takeoff clearance reaction time was defined as the time from ATC issuance of the 'Cleared for Takeoff' directive to the point the flight begins its takeoff roll. This time, as well as the takeoff roll duration, was not part of the surface system's prediction. Thus, an estimate was required prior to communicating the predicted wheels OFF time to the downstream decision support system in PDRC.

The clearance reaction time measure required undelayed ATC voice clearance instructions as well as direct line of site to the departing flight to determine when it began its takeoff roll. A total of 108 flights were observed from the DFW Center tower during May-July 2012.

2. Uncertainty Results

The results of the takeoff clearance reaction time measure are depicted in Fig. 7. An interesting observation was that approximately 35% of flights did not stop on the runway threshold but rather continued directly into their takeoff roll. For these flights, the departure clearance was given before the runway hold line or immediately upon entering the runway. These flights are captured in the first bin of the histogram in Fig. 7.

The mean and median for the sample taken were both 6 seconds. The data sample also demonstrated some variance with a standard deviation of 5 seconds, a minimum of 0 and a maximum of 25 seconds.

Using the mean and median times for the entire sample do not give the best indication of how long it took for the time it took the pilot to react to the departure clearance given this includes flights that did not stop at all. For those flights that did stop on the runway threshold, the mean and median clearance reaction time was 9 seconds.

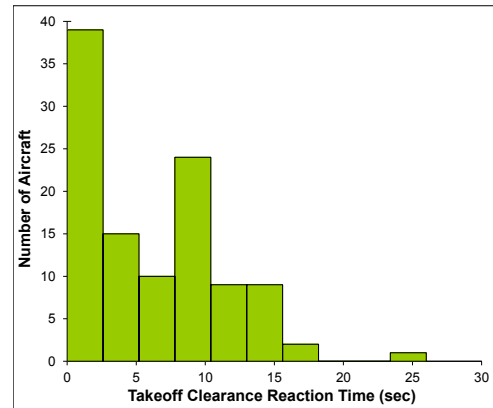


Figure 7. Scattergram showing duration from takeoff clearance to start of roll.

G. Takeoff roll duration

The purpose of the takeoff roll measure was to assess uncertainty that exists in the range of times from a flight's start of roll to the time at which the rear wheels were OFF the airport surface. An estimate of this measure was required given the downstream decision support system's ascent model begins at wheels OFF time and location. PDRC's SDSS component uses an adaptable value for takeoff roll duration. Currently, this value is the same for all aircraft types. The takeoff roll prediction error discussed in this paper is the difference in actual takeoff roll versus the mean takeoff roll.

1. Measurement Approach

Direct observation of takeoff duration was selected over surface data analysis techniques to ensure the accuracy of the start of roll and rear wheel liftoff. This measure required direct line of site to the departing flight in order to determine when the flight began its takeoff roll and wheels were off of the airport surface. One hundred ninety one (191) flights were observed from DFW center tower during 2011 and 2012 PDRC evaluations.

2. Uncertainty Results

The results of the takeoff roll duration measure are depicted in Fig. 8. Both the mean and median takeoff time duration for the sample taken was 38 seconds. Some variation was noted with a standard deviation of 7 seconds, a minimum of 18 seconds and a maximum of 55 seconds.

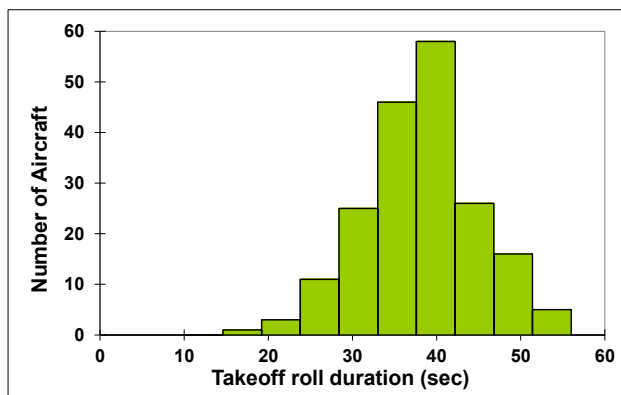


Figure 8. Takeoff roll duration.

Analytical models for calculating the distance for takeoff roll exist in the literature.⁵ These models generally require knowledge of takeoff weight, prevailing winds on the surface or other variables that are not currently available to the PDRC system in real time. Given this data access limitation, as well as a focus on takeoff duration rather than takeoff roll length, sampling of actual takeoff roll duration was utilized to formulate an average takeoff time duration. However, assuming a significant sample size it may be possible to narrow the variation by use of empirical data. In the sample of data collected for this research, the average takeoff duration for a McDonnell Douglas MD-80 aircraft was 41 second, compared to an average takeoff time for an Embraer ERJ 145 of 32 seconds.

H. TRACON transit time

The purpose of the TRACON transit time measure was to assess uncertainty that exists in the transit between the wheels off event and crossing the departure fix on the boundary of TRACON and Center airspace. The following paragraphs describe improvements to TRACON transit time predictions that were incorporated into PDRC's TMA/EDC component and used during the operational evaluation.

Currently, departure logic in PDRC's TMA/EDC component predicts that the flight will fly an adaptable number of nautical miles in the direction of departure and then acquire the first departure fix in the departure route. Analysis of the DFW departure data revealed that the first fix was significantly downstream in the aircraft's route of flight. Due to this, the en route transit time predicted by TMA/EDC assumed that the flight would head directly toward this fix instead of capturing the nominal waypoints along the RNAV departure route. Figure 9a illustrates the horizontal profile of the TMA/EDC predictions which are representative of the current operational system logic. In this diagram, DFW is the green dot and DARTZ (red X) is the first fix in TMA's estimated route. The thick, dull gray line is the actual track. The various colored lines which extend from the gray route are the TMA/EDC provided estimated routes at that point in time. In an ideal scenario, these lines would overlay the thick gray route.

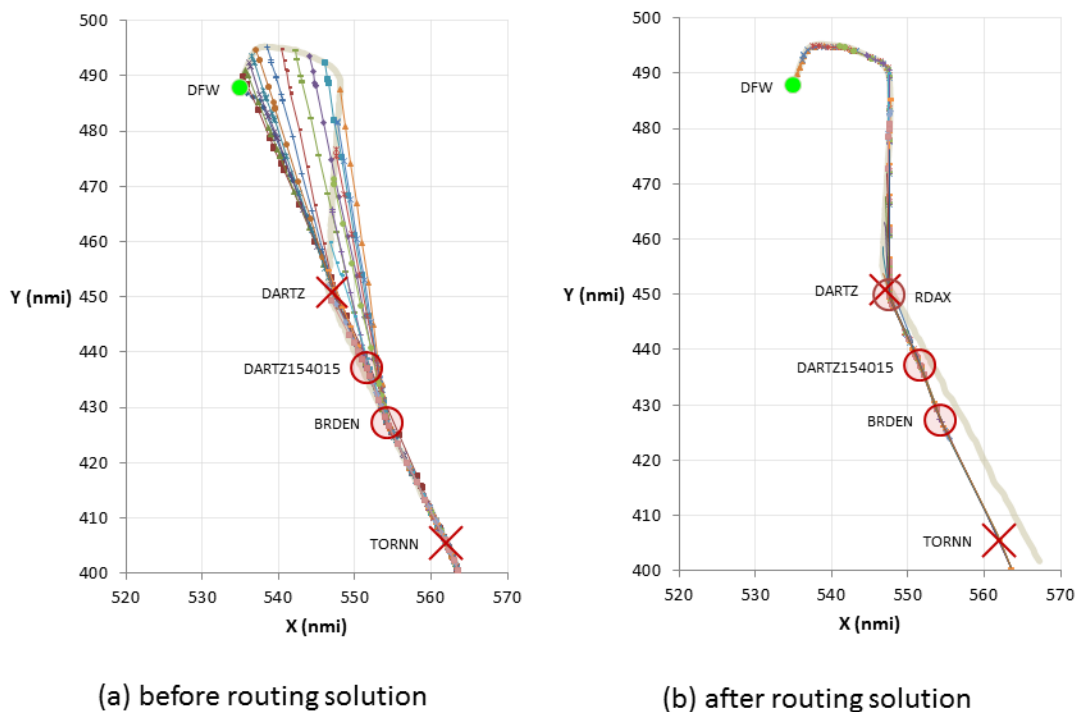


Figure 9. TMA departure route predictions before and after implementation of an adaptation-based solution.

In order to provide a more accurate route for PDRC evaluation, several potential solutions were analyzed. The solution selected was to create more specific departure routing which includes the expected TRACON departure fixes from the RNAV departure route. The routing assignment in adaptation was linked to the departure runway, which is automatically passed to TMA/EDC from PDRC's SDSS component. Figure 9b provides a graphical view of PDRC predictions of a DFW departure after implementing this solution. As illustrated the predictions and the actual tracks align closely.

The previous figures plot the route geometry before and after but do not reflect the impact to the estimate times-of-arrival to the departure fix. To determine this, the predicted and actual transit times were compared for a sample of 109 flights from DFW to IAH. Figure 10 plots the difference between the actual time of flight and the TMA predicted time of flight to the departure point using the routing change illustrated in Fig 9b. The actual transit time was defined as the time between the first radar track and the time when the flight crosses the departure fix. The data in this figure are stratified by departure runway. The points shown in blue diamonds were measures of TMA Estimated Time-of-Arrival (ETA) error to the meter fix before the routing solution was implemented. The same measure was taken for a sample of 53 flights after the routing solution previously mentioned was added to the system. The results are illustrated as green triangles in Fig. 10. North flow departures time prediction error to Runway 35 prior to the routing solution showed a mean error of 176 seconds, while after the solution this was reduced to a mean error to 28 seconds. Runway 36 demonstrated similar improvement with a mean error of 235 seconds prior to the routing solution that was reduced to 62 seconds after the solution.

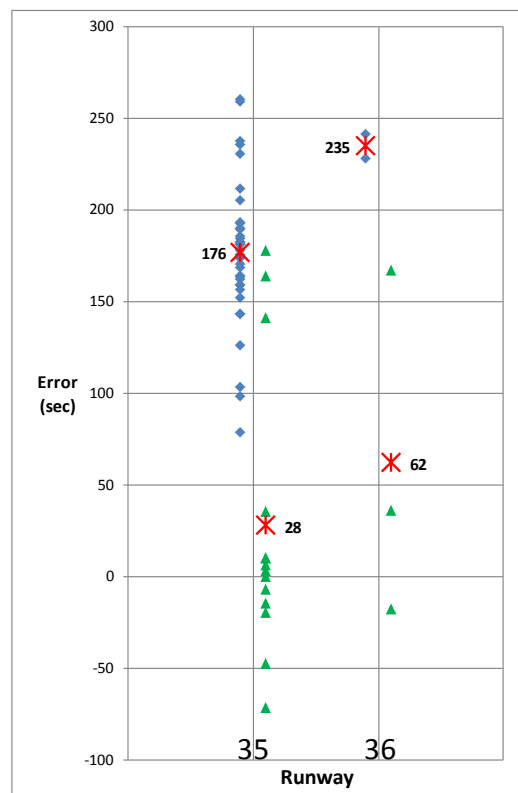


Figure 10. TMA ETA error before and after adaptation solution.

1. Measurement Approach

This measurement considered the transit between the wheels off event and crossing the departure fix on the boundary of TRACON and Center airspace. The truth data used to assess this uncertainty was the actual OFF time as derived from the surface system and actual departure fix crossing time derived from airborne surveillance from TRACON and Center data sources. TRACON transit time prediction error is any error associated with the PDRC system using the current day algorithms used in the field evaluation.

2. Uncertainty Results

The improvements mentioned in the previous section were incorporated into PDRC and used during the operational evaluation. Ninety two (92) flights from the PDRC evaluation were analyzed. Figure 11 provides a breakout of the size and frequency of TRACON transit time error. The transit time error is measured in absolute values to prevent aircraft that had negative flight time error (transit time lower than predicted) from biasing the results. For the PDRC scheduled flights during the operational evaluation, the mean absolute error was 25 seconds with a median TRACON transit time error of 21 seconds. Some variation was observed with a standard deviation of 20 seconds, a low error of 1 seconds and high of 122 seconds.

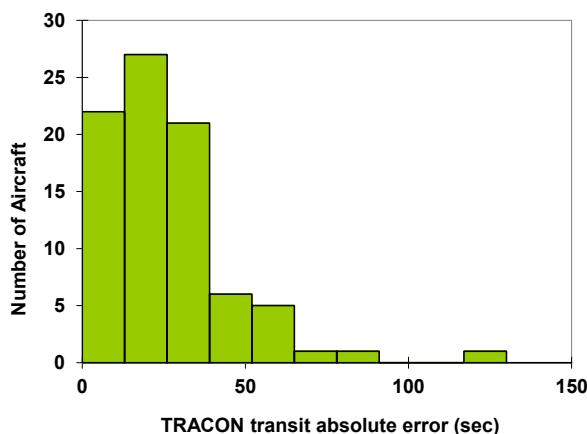


Figure 11. Absolute Error of TRACON Transit Time

The flights with the highest TRACON transit time error were scrutinized. Amongst these flights was an aircraft scheduled to depart on Runway 35L. This runway prediction was supplied by the surface system based upon statically adapted rules. Later, the flight changed to depart from Runway 36R. The runway the flight was scheduled with added approximately 40

seconds of additional transit time. This example highlights the need for obtaining the correct runway assignment prior to the en route scheduling process.

I. Center Transit Time

The purpose of the Center transit time measurement was to assess the uncertainty in transit between departure fix crossing and meter point crossing events.

1. Measurement Approach

The departure fix is located on the boundary between TRACON and Center airspace, while the departure meter points are generally located on the neighboring Center boundary. Truth data used to assess this uncertainty were derived departure fix crossing time and meter point crossing times from airborne surveillance. Those flights which had a change to the meter point assignment after scheduling were removed from the sample to eliminate flights which changed intent between the time the flight was scheduled and crossing of the meter point. Center transit time prediction error is any error associated with the PDRC system using the current day algorithms used in the field evaluation.

2. Uncertainty Results

The results of the Center transit time uncertainty measure are illustrated in Fig. 12. The mean Center transit time error for all PDRC scheduled flights was 49 seconds with a median error of 32 seconds.

Note that the mean error is approximately twice as high as the TRACON transit time error despite the fact that the flight distance for these two measures are approximately the same.

A significant factor is this error is that the PDRC evaluation used the EDC component of the TMA decision support tool which did not present times to the sector controllers' scopes. Thus, sector controllers made sequencing choices independent of the PDRC guidance which introduces individual sequencing preferences into the uncertainty.

Manual observation of PDRC scheduled flights and discussions with Center Traffic Managers also revealed other factors, including significant speed fluctuations in the overhead stream, flights that cut corners off of the nominal route, pop-up flights that were scheduled after the PDRC scheduled flights, and altitude error. In the case of altitude error, the primary challenge was that the TMA/EDC system has no knowledge of the Letter of Agreement between Fort Worth Center and Houston Center in which aircraft are provided to Houston at flight level 290 if Houston is in East flow and flight level 310 if Houston is in West flow. Without the crossing altitude information TMA/EDC is left to speculate that the flight will cross at their filed flight plan altitudes which could have significant differences in wind speed and/or could take some time to maneuver to.

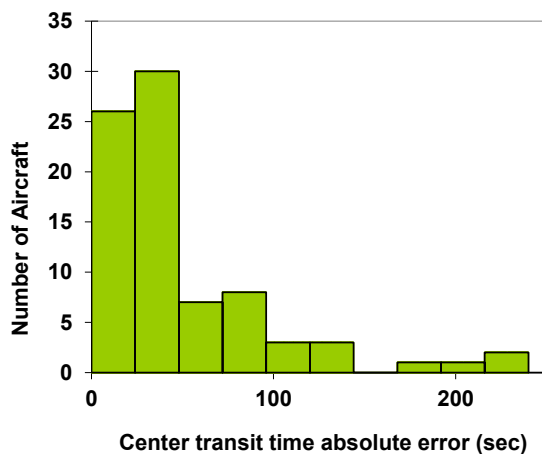


Figure 12. Absolute Center transit time error.

III. System Performance

To determine system performance, PDRC research utilized a combination of quantitative metrics and qualitative feedback from operational personnel and subject matter experts. The focus of this section is on the objective metrics that were used to assess PDRC system performance.

A. OFF Time Compliance

One objective of PDRC is to improve upon schedule compliance by reducing uncertainty that has been demonstrated in manual coordination.³ Thus, OFF time compliance is an important system metric for PDRC.

1. Measurement Approach

The approach used in this measurement was to leverage the highest precision measurement source available to evaluate compliance. For the baseline sample, a full year of OFF time compliance data were available from operational TMA/EDC recordings covering more than 400 scheduled flights from October 2010 until November 2011. Flights with strategic times (EDCTs) were removed from the compliance analysis presented here but utilized in other analyses. EDCTs were not counted in the primary compliance measure because they introduced variation due to procedural differences which were not the focus of this research. Briefly, the research team observed individual controllers following different procedures in situations where flights were subject to both a strategic and tactical TMI. The OFF time agreed upon between Center and Tower traffic managers was defined to be the coordinated OFF time, which was compared against the departure time as obtained from the departure message from en route automation. Given that this large sample of data covered a long duration in which unknown circumstances might have been involved without a PDRC observer to report them, the outliers outside of 1.5 times the IQR were removed.

Measuring PDRC OFF time compliance was more straightforward than the baseline given firsthand knowledge of every scheduled flight. For example, one PDRC scheduled flight that was subject to an APREQ procedure was later expedited in order to prevent potential hail damage. At the point that verbal direction was given to expedite the flight, the APREQ time was no longer valid. However, no electronic commands were issued for this flight and had the team not been aware of this occurrence then the flight would otherwise have looked non-compliant. Flights which had both a strategic and tactical TMI were captured for analysis in PDRC, but like the baseline set they did not count toward OFF time compliance results. The measure of OFF time compliance for a PDRC flight is the coordinated OFF time versus the actual wheels off time as available in the PDRC. PDRC calculates the actual OFF by utilizing a detected start of takeoff roll and adapted roll duration.

A total of 120 flights were scheduled by the PDRC system during the operational evaluation from May 30, 2012 through July 26, 2012. For a flight to count as a PDRC scheduled flight, both the surface and the Center traffic managers had to schedule the flight using the PDRC system and agreed upon scheduling procedure.

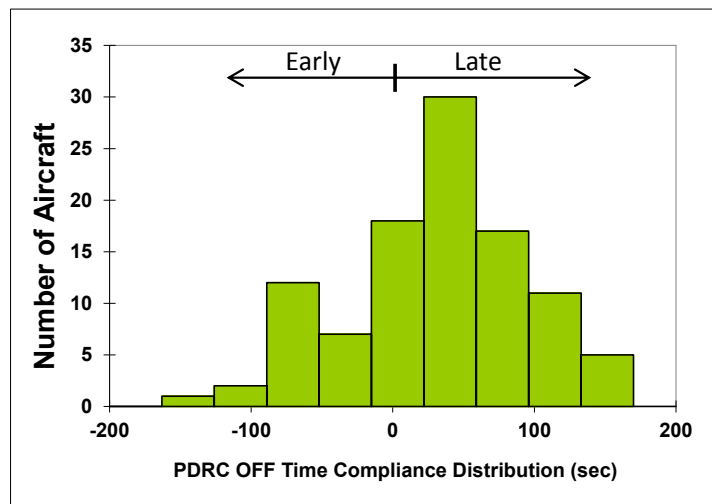


Figure 13. PDRC OFF time compliance distribution.

2. Results

The distribution of PDRC OFF time compliance is illustrated in Fig. 13. The mean compliance was 33 seconds with a median of 37 seconds, indicating a slightly later actual OFF time than planned time on average. A fair amount of variance was exhibited in this sample as well, with a standard deviation of 63 seconds, a minimum of 135 seconds early and a maximum of 165 seconds late. The two flights with the highest OFF time error were both due to the flight not having its weight and balance numbers when it arrived at the runway threshold.

Table 1 provides a comparison between the baseline OFF time compliance and the PDRC system OFF time compliance. The second column of this table has the PDRC OFF time compliance mean, median and standard deviation values. The baseline OFF time compliance is listed in the third column, and the last column indicates the estimated percentage of PDRC OFF time compliance compared to the baseline compliance. This improvement is characterized as a lower bound estimate due to the fact that outliers were removed from the baseline data set but not the PDRC data set. As previously discussed, outliers were removed given the PDRC team did not observe all flights in the one year sample. As the table indicates, PDRC scheduled flights demonstrated a significant improvement over baseline levels of OFF time compliance.

Table 1. PDRC absolute OFF time compliance compared with baseline.

	PDRC absolute OFF time compliance (sec)	Baseline OFF time compliance, outliers removed (sec)	OFF time error versus baseline (lower bound %)
Mean	59	108	55
Median	54	79	68
Standard Deviation	40	96	42

B. Hit Slot Performance Measure

The purpose of the hit slot performance measure was another way to assess how well the system delivered flights to the available airspace, or slot, they were originally scheduling into.

1. Approach

The purpose of the hit slot performance measure was to assess how well the system delivered flights to the available airspace, or slot, they were originally scheduling into. The hit slot measure compares the leading and trailing flight at the time of scheduling with those at the time the flight crosses the meter point in en route airspace. A detailed explanation of this measurement and the baseline hit slot results can be found in prior PDRC research.¹

2. Results

Forty-one percent (41%) of PDRC scheduled flights hit the slot they were scheduled into. This represents only a modest improvement over the baseline level of hit slot performance of 39%.¹ The merge of PDRC scheduled flights into the overhead stream was analyzed to determine the primary reason for low hit slot performance. The most common cause for missing the scheduled slot was a change in the overhead stream from the time the flight was scheduled to the time the flight arrived at the meter point. Departure uncertainty from nearby airports as well as airborne flights that short cut their route contributed to the overhead stream uncertainty.

IV. Discussion

This section discusses inferences of system performance from individual departure predictive accuracy measures.

A. Impact of Communication Uncertainty and APREQ Window Size on OFF Time compliance

Table 1 of this paper indicated PDRC exhibited a significant improvement over baseline absolute OFF time compliance. The primary reasons for improved OFF time compliance using PDRC is inferred to be reduced communication uncertainty between the Tower and Center traffic managers and use of a PDRC-enabled target time rather than the standard APREQ window. This section provides a summary of information gained during the PDRC evaluation from direct observations and traffic management controller (TMC) interviews.

Without the use of PDRC, the coordinated OFF time window for an APREQ is communicated verbally over facility inter-phone from the Center to the Tower. Currently, there is no set national standard for the time window to use in this communication or the phraseology to employ in this procedure. However, this time window is generally accepted as being 3 minutes and is structured to favor flights departing early rather than late. The general rule is to build a time window two minutes ahead (-2) and one minute behind (+1) the flight's desired OFF time. Observations of this verbal exchange and interviews of Center personnel indicate a significant degree of uncertainty in this communication. For example, a time of 17:25:26 Zulu from the TMA/EDC system is truncated and displayed in minutes level granularity to Center personnel as 1725. The Center TMC then verbally communicates the APREQ window which may be communicating a single value of 1725 or any number of variations in APREQ window size, bias and phraseology.

After receiving the time window verbally from the Center TMU, the Tower TMC must interpret the information. In some cases, the TMC receiving the information may assume the beginning of the first minute to the end of the last minute given. In the case of the time communicated as 1724 to 1727, this interpretation would be 1724:00 to 1727:59, which would be a 3 minute and 59 second window. Other Tower TMCs indicate that they take the specified window to mean the beginning of the first minute to the beginning of the last minute given.

Interviews with Tower personal revealed that another source of communication uncertainty is the location at which the flight is expected to be at this time window. In some cases, the Tower TMC assumed the location that the flight was expected at the negotiated time was when the flight was "tagged up", or the point at which TRACON

surveillance was first received for the flight. However, the trajectory from the TMA/EDC system begins at wheels OFF from the airport surface. Observations of “tag up” as compared to actual OFF time indicated this duration has a 26 second mean with 25 second median.

To reduce the complexity of this communication, the PDRC system automatically sent the time expected by en route automation at seconds level precision to the surface system. The Tower TMC then used the seconds level precision to communicate a minute’s level precision value to the surface local controller. The local controller was asked to try to achieve wheels OFF at the exact minute communicated to the best of their ability rather than using the standard APREQ compliance window described above.

B. Combined View of Departure Error

Earlier sections of this document described departure event measurements of uncertainty for departure events analyzed in this research. However, a reasonable question may be, how well does the system perform in the presence of this uncertainty? To answer this question, it was necessary to analyze the level of predictive error associated with each departure event. In most cases, the PDRC prediction for the event was utilized to obtain this measure. In some cases, like in the hypothetical case of scheduling flights from the gate which was not performed during the PDRC evaluation, it was necessary to estimate the level of error using a reasonable approach at the prediction like those discussed in earlier sections of this paper.

Figure 14 combines the departure prediction measures described in this research into a single diagram. This figure provides a view of the size and distribution of the current prediction error for each departure event in the presence of current day uncertainty. In the majority of cases, the PDRC prediction for the event is utilized to obtain this measure. In some cases, like in the hypothetical case of scheduling flights from the gate which was not performed during the PDRC evaluation, it was necessary to estimate the level of error using a reasonable approach at the prediction. The mean values of each measure are shown in bold near the red cross. The median value, upper quartile and lower quartile define the boundaries of the box structure in the box plot, and the ‘whiskers’ extend on both directions of the box to encompass the variance of the distribution without inclusion of the outliers. In the case of pushback start and center transit error, a portion of the box plot whisker cannot be seen because they go beyond the scale of the diagram.

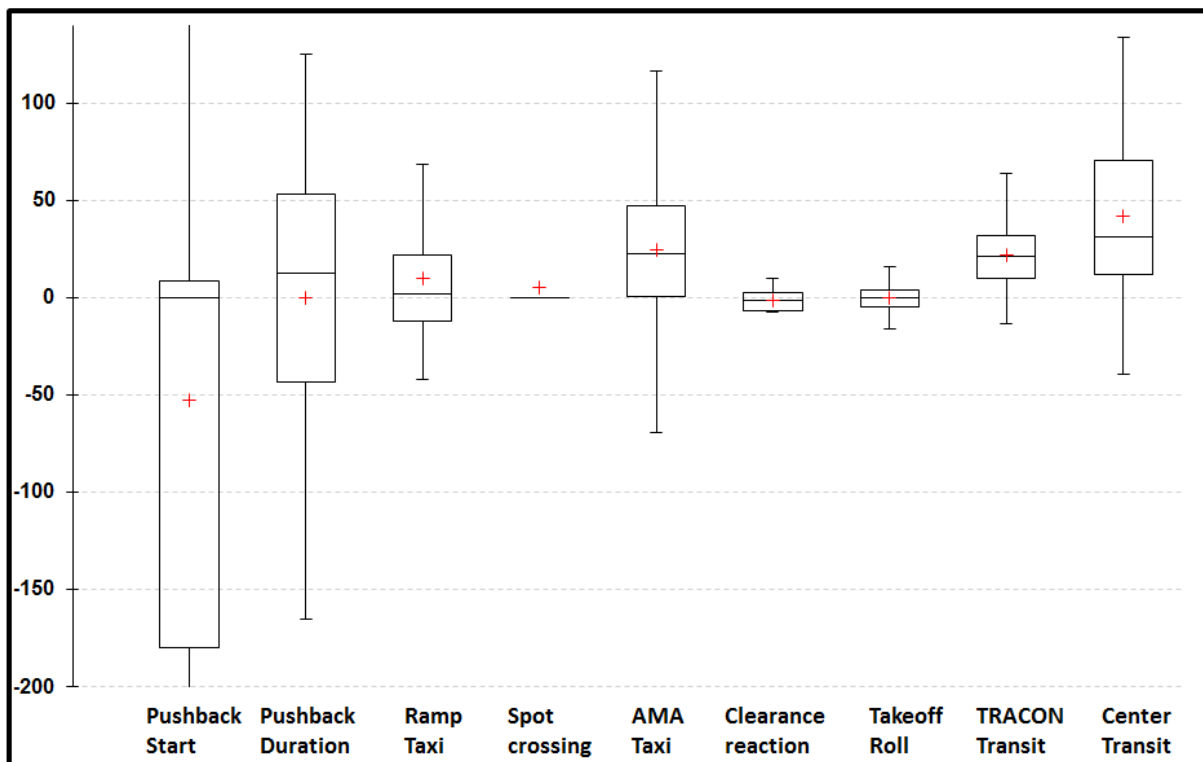


Figure 14. Comparison of prediction error for the various departure events analyzed in the PDRC operational evaluation.

The variance in the pushback start error, pushback duration error, AMA taxi error and Center transit error are the largest among all departure events. Of those events, only the AMA taxi and the Center transit error are part of current day tactical departure scheduling. Both the pushback duration and pushback start error would need to be considered when performing tactical scheduling prior to the spot.

Even assuming perfect OUT time compliance, the pushback duration event alone experiences variances that are high enough to prevent a flight from being able to make the current day tactical departure window which only allows flights to be one minute late. Additionally, while mean ramp taxi error is low based upon the average ramp taxi speed method utilized, the upper tail stretches nearly 70 seconds. If significant ramp taxi error occurred then this would leave little room for error for any remaining event predictions. However, the error measured in this analysis is absent air carrier efforts to meet a specified time. It is possible that with greater air carrier involvement flights could be expedite to meet an earlier time if required.

Spot crossing duration has a positive overall mean of 6 seconds. While this value is low, the maximum spot crossing error was 100 seconds and spot crossing duration is an event that can add to overall system uncertainty in current day tactical departure scheduling.

The AMA taxi represents a significant source of uncertainty in today's tactical departure scheduling process. The taxi time error has a number of error components, the most sizable of which are runway prediction error from the spot and predicted taxi speed from the spot.

The size of the error of the clearance and the takeoff roll time are negligible in comparison with the other error sources once taken into account. If they were not taken into account however, then on average each flight would have 6 seconds of clearance error plus 38 seconds of roll time error, for a total of 44 seconds error on average. While this seems small, that equates to approximately 5nmi in the overhead stream (assuming 420 kts).

TRACON transit time error has been reduced significantly with the use of more detailed TRACON routing in TMA/EDC as well as automatic utilization of the runway assignments passed from the surface system. However, the overall average of the error is positive. That is, the current system is consistently under predicting the transit time from OFF to the departure fix. For tactical scheduling purposes, it is better to over predict rather than under predict TRACON transit time if forced to choose between the two. Over predicting the transit time would allow the flight to be delayed to meet the time rather than accelerated.

Center transit time has the largest positive error of any departure event measured. Bearing in mind that all the flights measured were during a volatile period in which an APREQ event was put in place to help manage, a certain degree of variance is not surprising. Additionally, this operational evaluation involved only outbound tactical departure scheduling³ using the TMA/EDC decision support tool. Currently, the FAA operates TMA/EDC in an open-loop mode. Unlike arrival metering with TMA, TMA/EDC schedule times and sequence information are not displayed on sector controllers radar scopes. Center TMCs use TMA/EDC to manage constrained traffic flows to provide sector controllers with a workable traffic situation. Sector controllers solve the traffic puzzle with no knowledge of the TMA/EDC planned solution. Thus, differences between the TMA/EDC and controller solutions are to be expected. Given the timing associated with widespread deployment of a surface capability that could supply the OFF times required, it is likely better to assume a metering environment in which times are presented to the controllers with +/-30 seconds of error like those demonstrated by the Efficient Descent Advisor (EDA).⁶

C. Extending tactical departure scheduling to the gate

The cumulative departure scheduling error that may occur when scheduling from the gate was estimated by taking 1000 random draws, with replacement, from the error data samples described in previous sections of this paper. The error components of each departure event were accumulated to form the total surface departure error associated with each of the 1000 flights in the sample. If the total departure error was one minute or less, the flight was considered compliant with the surface OFF time. Total departure error that allowed a flight to be available earlier than their scheduled time was considered compliant for this measure given the flight could be delayed by ATC to meet the required time.

The cumulative error estimate for PDRC during the operational evaluation was taken by using the takeoff roll time and clearance reaction time that were available during the evaluation, given these values were slightly different than those discussed in this research. As indicated in Table 2, scheduling from the spot using the levels of PDRC accuracy available during the PDRC operational evaluation yielded an estimated 71% departure compliance. That is, 71% of flights had one minute or less total surface error when scheduling from the spot. For comparison purposes, the actual percentage of flights in the PDRC operational evaluation with one minute or less surface error were determined to be 70%. Thus, the theoretical estimates and actual distribution for PDRC OFF compliance are very close.

Table 2. Estimated OFF time compliance percentage using cumulative departure event error

Scheduling from the spot using PDRC accuracy during operational evaluation	Scheduling from the spot using changes described in this research	Scheduling from the gate using PDRC accuracy during operational evaluation	Scheduling from the gate with changes described in this research
71%	92%	64%	73%

An estimate was taken of the improvement to OFF time compliance that could be achieved from implementing changes described in this research. Specifically, the prediction error associated with using the mean clearance reaction time, mean takeoff roll time and adding a 30 seconds buffer to remove late bias associated with current OFF time compliance. The estimated OFF time compliance when scheduling from the spot with these changes is 92%.

OFF time compliance while scheduling from the gate was estimated by using the error components associated with pushback start, pushback duration and ramp taxi. Scheduling from the gate using the levels of PDRC accuracy available during the operational evaluation yielded an estimated 64% compliance. However, scheduling from the gate with the changes described in this research would yield an estimated 73% OFF time compliance. For comparison, OFF time compliance achieved nation-wide with manual scheduling process is approximately 69%.¹

It is important to note that this analysis did not consider active air carrier participation in the tactical departure scheduling process. Currently, the PDRC concept places no requirements on air carriers beyond passively providing gate assignment and pushback estimate information that already resides in air carrier systems. Other concepts such as Spot and Runway Departure Advisor (SARDA),^{7,8} Collaborative Departure Queue Management (CDQM),⁹ and Surface Collaborative Decision Making (Surface CDM)¹⁰ assume an active air carrier role in departure scheduling to enable surface delays to be absorbed at the gate or in the ramp area. Combining active air carrier participation from these concepts with PDRC's integration between Tower and Center departure scheduling systems may enable the tactical departure scheduling horizon to be extended to the gate with satisfactory compliance.

D. Removing Late Bias from Departure Fix compliance

Given only modest improvements to hit slot performance over the baseline measure despite PDRC's significant improvement to OFF time compliance, an alternative measure of system performance was developed. The purpose of the departure fix compliance performance measure was to assess how well the system delivered flights to the boundary between TRACON and Center airspace compared to the scheduled time. This measure was useful because it provided an interim point between wheels off and meter point crossing at which the performance of the system could be analyzed. In addition, the majority of airborne vectoring and speed controls occur after the departure fix which allows an objective metric to be obtained from operational data with fewer confounding influences than the hit slot measure.

To determine if the departure fix compliance was a reliable performance measure, PDRC OFF time compliance and departure fix crossing errors were analyzed for correlation. The correlation coefficient for these two data sets was 0.937, which indicates a high correlation. This suggests that as OFF time uncertainty decreases, so does departure fix crossing uncertainty. In contrast, the correlation coefficient between PDRC OFF time compliance and meter fix compliance was only 0.16, indicating a low correlation.

The departure fix compliance measure utilized in PDRC relied upon the coordinated departure time negotiated between the systems and TMA/EDC's estimate of time to fly to the departure fix. The two values were combined to form the scheduled departure fix crossing time. The scheduled crossing time was then compared against the actual crossing time to determine if it was compliant. The compliance standard used was the same as today's standard of two minutes early through one minute late (written as -120/+60). The results for departure fix compliance for PDRC scheduled flights are that 51% hit the -120/+60 second window.

Given significant improvement to OFF time compliance demonstrated with PDRC as well as lower variance in the overall distribution, one might expect higher departure fix compliance than was demonstrated. Analysis of departure fix compliance revealed that the primary reason for lower than expected departure fix compliance is flights were slightly later than planned. This fact is not surprising given that the local controllers using the PDRC times were required to change their procedure to use a single minute time rather than the time window that they were accustomed to. To compensate for this, a slight modification can be made to the times that are communicated via PDRC. This modification can be implemented in the software communication layer between the Center system

and the Tower system such that no changes to PDRC scheduling procedures would be required. Analysis of possible buffer values using PDRC scheduled flights indicated departure fix compliance as high as 86% may be possible with this simple software change.

V. Conclusions

During the operational evaluation of Precision Departure Release Capability at DFW, OFF time compliance improved from an average absolute error of 108 seconds to less than 59 seconds. PDRC demonstrated greater predictability than the baseline sample by decreasing OFF time error from a standard deviation of 96 seconds to 40 seconds.

Significant improvements to TRACON transit time predictions were achieved by including TRACON-specific routing in the horizontal profile and electronically supplying the airborne system with the departure runway assignment from the airport surface system. Despite these improvements, additional work is needed to reduce TRACON transit time error.

Prediction errors associated with the departure events were utilized to estimate the OFF time compliance associated with scheduling tactical departures from the gate. This estimate indicates that OFF time compliance of 73% of flights scheduled from the gate may be possible without requiring active airline involvement, which exceeds baseline tactical departure scheduling OFF time compliance.

Acknowledgements

The authors would like to acknowledge the essential support provided by FAA personnel at the Fort Worth Center Traffic Management Unit and Dallas/Fort Worth ATCT. Finally, we wish to thank our colleagues at NTX and NASA Ames whose support was critical to the success of PDRC research objectives.

References

¹Capps, A. and Engelland, S.A., "Characterization of Tactical Departure Scheduling in the National Airspace System," AIAA-2011-6835, *11th American Institute of Aeronautics and Astronautics (AIAA) Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, 20-22 Sep. 2011.

²Grabbe, S., "Traffic Management Advisor Flow Programs: an Atlanta Case Study", *American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation, and Control Conference*, Portland, Oregon, 8-11 Aug. 2011.

³Engelland, S.A. and Capps, A., "Trajectory-Based Takeoff Time Predictions Applied to Tactical Departure Scheduling: Concept Description, System Design, and Initial Observations," AIAA-2011-6875, *11th American Institute of Aeronautics and Astronautics (AIAA) Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, 20-22 Sep. 2011.

⁴Palopo, K., Chatterji, G., and Lee, H., "Benefit Assessment of the Precision Departure Release Capability Concept," AIAA-2011-6834, *11th American Institute of Aeronautics and Astronautics (AIAA) Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, 20-22 Sep. 2011.

⁵Middleton, D.B., Srivatsan, R., and Person, L.H., Jr., "Flight Test of Takeoff Performance Monitoring System," NASA TP-3403, 1994.

⁶Coppenbarger, R., Dyer, G., Hayashi, M., Lanier, R., Stell, L., Sweet, D., "Development and Testing of Automation for Efficient Arrivals in Constrained Airspace," *27th International Congress of the Aeronautical Sciences (ICAS)*, Nice, France, 19-24 Sep. 2010.

⁷Jung, Y. C., Hoang, T., Montoya, J., Gupta, G., Malik, W., and Tobias, L., "A Concept and Implementation of Optimized Operations of Airport Surface Traffic," *10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Fort Worth, TX, 13-15 Sep. 2010.

⁸Hoang, T., Jung, Y., Holbrook, J., and Malik, W., "Tower Controllers' Assessment of the Spot and Runway Departure Advisor (SARDA) Concept," *9th USA/Europe ATM R&D Seminar (ATM2011)*, Berlin, Germany, 14-17 June 2011.

⁹Brinton, C., Lent, S., and Provan, C., "Field Test Results of Collaborative Departure Queue Management," *29th Digital Avionics Systems Conference*, Salt Lake City, Utah, 3-7 Oct. 2010.

¹⁰FAA Air Traffic Organization Surface Operations Office, "U.S. Airport Surface Collaborative Decision Making (CDM) Concept of Operations (ConOps) in the Near-Term – Application of Surface CDM at United States Airports," 15 June 2012