Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)

Exploring Wind Information Requirements for Four Dimensional Trajectory-Based Operations*

Tom G. Reynolds+, Michael McPartland+, Tom Teller# & Seth Troxel+
+Air Traffic Control Systems Group & #Surveillance Systems Group
MIT Lincoln Laboratory
Lexington MA, USA

***Abstract*—Many future air traffic control concepts depend on access to high accuracy wind data due to time-based control elements, such as required time of arrival at a meter fix under 4D-Trajectory-Based Operations. Errors in the wind information relative to the truth winds could significantly degrade the performance of the procedure. Unacceptable performance could be mitigated by improving wind information accuracy by using higher accuracy forecast models, updating wind information more frequently, or upgrading the way winds are handled in the avionics systems. This paper: (1) establishes the relationship of wind information accuracy to 4D-TBO performance for a selection of operationally relevant scenarios to identify wind needs to support them, and (2) presents examples of how this information can be used to determine what wind information content and update rate to the aircraft will deliver a given target performance level to help inform concept of operations development and supporting technology needs.**

***Keywords*—four dimensional trajectory-based operations, wind requirements, Required Time of Arrival, flight management system.**

I. INTRODUCTION

Many NextGen applications depend on access to high accuracy wind data due to time-based control elements, such as Required Time of Arrival (RTA) at a meter fix under 4D-Trajectory Based Operations (4D-TBO) (also referred to as Controlled Time of Arrival (CTA) and Time of Arrival Control (TOAC) respectively) procedures. Figure 1 illustrates how wind information is used on the ground by Air Traffic Control (ATC) to develop time targets for use in a 4D-TBO procedure and for flight planning by airlines, and then wind information in the aircraft is used by the avionics to manage the aircraft trajectory to these targets. The performance of the procedure is typically measured as a mean and 95% spread of RTA compliance error at the meter fix. Target performance is likely to be specified as a maximum allowable RTA compliance error a given fraction of the time, for example +/- *x* seconds 95% of the time. Any errors in the ground and/or aircraft wind information relative to the truth winds actually flown through can significantly degrade the performance of the procedure, resulting in a wider RTA compliance error distribution at the meter fix. Unacceptable performance could be mitigated by improving wind information in the aircraft, for example by using higher accuracy wind forecast models to generate wind inputs for the ground or airborne systems, updating wind information more frequently, or improving the way winds are handled in the avionics systems.

Prior work has established the important impact that wind errors can have on the ability to meet a given RTA target, e.g., [1]. For example, flight trials in Europe [2] found that the accuracy of the forecast wind used by the Flight Management System (FMS) of the aircraft significantly affected the RTA performance, especially in later stages of the descent where the potential to make sufficient corrective speed changes is severely limited. Similarly, trials in the US [3] found the need to update FMS winds shortly before top of descent in order to increase the fraction of flights which could complete the intended RTA procedures. Other studies have taken a more focused look at the impact of wind forecast errors on FMS calculations. For example, [4] focused on the effect of wind forecast errors on open-loop Estimated Time of Arrival calculations. The study reported in this paper builds upon this prior work by focusing on wind information quality impacts on closed-loop RTA performance, and then illustrating how this information can be used to guide the development of procedures and associated technologies.

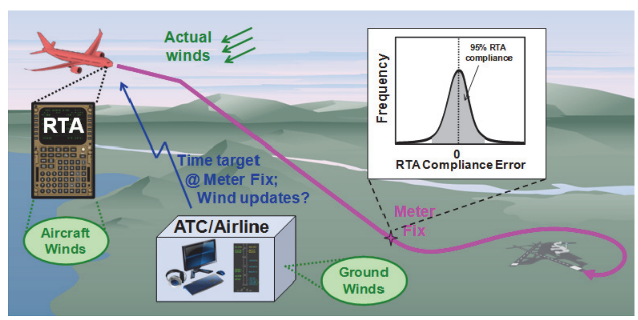

Figure 1. Wind Information in 4D-TBO Procedures

II. WIND INFORMATION ANALYSIS & IMPLICATIONS ASSESSMENT

The objectives of this study are to: (1) Establish the relationship of wind information accuracy to 4D-TBO performance for a selection of operationally-relevant scenarios to identify wind needs to support them; and (2) Present

*This work was sponsored by the Federal Aviation Administration (FAA) under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

examples of how this information can be used to determine what wind information content and update rate to the aircraft will deliver a given target performance level to help inform Concept of Operations (ConOps) development and supporting technology needs. In order to address objective (1) to explore the relationship of wind information to NextGen application performance, a Wind Information Analysis Framework has been developed [5,6] as shown in Figure 2.

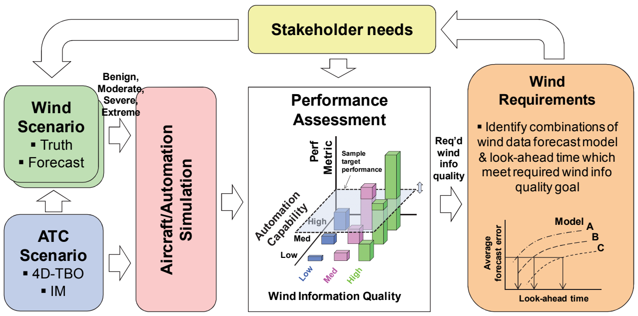

Figure 2. Wind Information Analysis Framework

In the framework, the *ATC Scenario* represents the characteristics of the ATC environments for the application of interest, e.g., specifics of the procedures, infrastructure, demand levels, equipage, etc. The *Wind Scenario* element represents the "truth" wind environments of relevance to the ATC scenario being studied (hence the arrow from the ATC Scenario to the Wind Scenario block), as well as the characteristics of different "forecast" winds relative to the actual wind field experienced. Truth wind fields are selected to expose the aircraft to various representative conditions to test response across a range of operationally-realistic situations. In addition to wind speed and direction, the wind scenario data include associated atmospheric variables needed to accurately simulate aircraft performance including temperature, pressure, and humidity. The *Aircraft/Automation Simulation* element represents the behavior of the aircraft, engine, autopilot and FMS in the context of the wind and ATC scenario being studied. By running simulations of how aircraft perform in the context of a given ATC application with varying qualities of wind forecasts while flying through various truth wind fields, it is possible to build up a trade-space of performance as a function of key independent variables such as wind information quality and aircraft automation capability. This is illustrated in the *Performance Assessment* element of the framework. This trade-space can then be used to establish what level of performance may be expected from a given wind information quality and automation capability combination. If a certain level of performance is required, this would define a horizontal slice through the trade-space from which combinations of wind information quality and aircraft automation capability that exceed that standard can be identified. The *Wind Requirements* element identifies which combinations of wind data content, wind forecast model and look-ahead times (i.e., the difference between the forecast issue time and its valid time) meet the wind information quality level identified from the previous element that achieve the target performance level. Finally, the *Stakeholder Needs* element represents the key role of stakeholders in determining appropriate choices in the other framework elements, e.g., in terms of which scenarios and performance metrics are of value to them to support the creation of guidance or requirements documents or to inform appropriate Concept of Operations (ConOps). The key stakeholders consulted on this work were a range of Radio Technical Commission for Aeronautics (RTCA) Special Committees (SC-206, 214, 227 and 186) with representation across the FAA, airlines and industry.

The framework is designed to be scalable with respect to scope and fidelity of its individual elements, as well as flexible with respect to the specific ATC application being studied. In Phase 1 of this project, the utility of this framework was initially demonstrated using a simplified version of the framework elements applied to a single 4D-TBO scenario. This demonstrated significant insights that could be generated from its use, as discussed in [5]. Phase 2 and 3 further refined the 4D-TBO analysis and expanded into Interval Management applications [6].

In order to address objective (2) to interpret the findings from the Wind Information Analysis Framework, a complementary Wind Information Implications Flow Diagram has also been developed as shown in Figure 3 which comprises six steps:

1. Define the scenario of interest, corresponding to an ATC Scenario case previously assessed with the Wind Information Analysis Framework.
2. Identify the appropriate Wind Information Analysis Framework performance trade-space corresponding to the scenario of interest from step 1.
3. Select a target performance level desired to be achieved by aircraft within the context of the scenario of interest.
4. Establish whether feasible combinations of performance drivers meet the target performance level selected in step 3, i.e., are there combinations of wind information quality and automation capability which meet/exceed the desired performance level
   a. If "YES", then identify required wind information level and proceed to step 5.
   b. If "NO", then a need has been identified to either
      i. Select a lower target performance level (go to step 3), or
      ii. Develop enhanced wind models or automation capabilities for analysis with the Wind Information Analysis Framework, resulting in a new trade-space (go to step 2).
5. Establish whether feasible combinations of wind forecast models and look-ahead times exist to meet/exceed the required wind error limit established from step 4.
   a. If "YES", then proceed to step 6
   b. If "NO", then a need has been identified to either
      i. Select a lower target performance level (go to step 3), or
      ii. Develop enhanced wind models or automation capabilities for analysis with the Wind Information Analysis Framework, resulting in a new trade-space (go to step 2).
6. Identify procedure, ConOps and supporting technology needs to get wind information of the required accuracy to the aircraft and ground systems for operational use that support the required performance for the scenarios of interest.

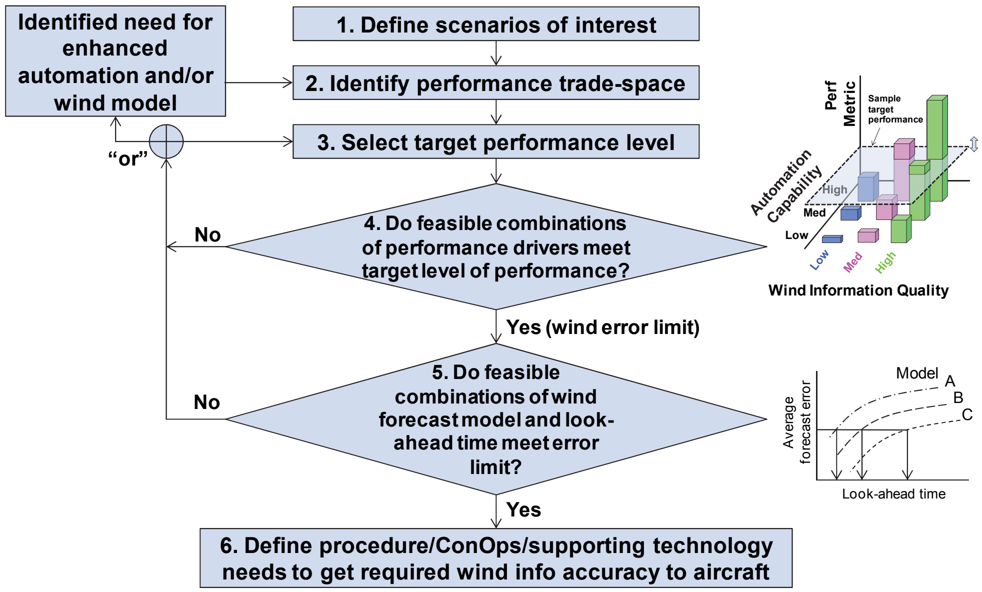

Figure 3. Wind Information Implications Flow Diagram

The next sections discuss the analysis which develops the trade-spaces for use in step 2 and the wind forecast model performance for use in step 5. The final section provides an example of how each are used in the context of this diagram for a case study application.

III. Analysis of Wind Information Accuracy on 4D-TBO Performance

This section details analysis which has been conducted to quantify the performance of a variety of 4D-TBO procedures under a range of truth and forecast wind conditions to better understand the impact of wind information quality. Two key performance metrics are used in the analysis in this section:

- RTA Time Error (RTA TE): The difference (in seconds) between the actual time achieved at the meter fix and the RTA time target assigned. Positive values indicate that a flight crossed the meter fix late, negative values indicate a flight crossed early.
- RTA Time 95% Confidence Interval (RTA 95%CI): The mean and estimated interval containing 95% of RTA TE for the particular conditions in question, calculated as $\mu \pm 2\sigma$ (Mean±2×Standard Deviation assuming Gaussian distribution) of the distribution of time errors.

The analysis in this section builds upon baseline data from a 2011 flight demonstration of 4D-TBO procedures and reports data from simulations conducted using multiple types of FMS systems in varied wind and ATC scenarios.

A. Baseline 2011 Seattle Flight Trials

Live flight trials were conducted under another program which provide high value inputs to this analysis. The trials occurred in November/December 2011 involving Alaska Airlines (ASA) and FAA's Seattle ARTCC (ZSE) [7]. In these trials, RTA operations were conducted during arrivals via the three most heavily utilized corner posts (arrival fixes) at Seattle-Tacoma International Airport (SEA). All ASA B737 flights equipped with General Electric (GE) FMS systems and arriving to SEA during operational hours via the specified procedures were candidates for participation. About 70 minutes prior to arrival, candidate aircraft received and loaded the Rapid Update Cycle (RUC) wind forecast applicable to the estimated time of arrival for their arrival route, via the Aircraft Communications Addressing and Reporting System (ACARS). Candidate aircraft then downlinked intent data including the range of achievable times at the meter fix to the Traffic Management Unit (TMU) at ZSE. The TMU determined and posted the RTA time according to the Traffic Management Advisor (TMA) scheduling tool. A total of 833 flights were assigned an RTA to the meter fix by ATC during the 23 days of flight trials, of which 595 flights (71%) fully executed and completed the RTA. Of the flights completing the assigned RTA, 86.4% crossed within 20 seconds of the RTA and 96.6% crossed within 30 seconds of the assigned time. Nearly all flights completing RTAs met the altitude/speed crossing restrictions (unless otherwise instructed/cleared by ATC). A summary of RTA performance in this demonstration is depicted in Figure 4. The mean RTA TE for the flights completing RTAs in all conditions was 9.0 seconds late, while the estimated RTA 95%CI was -13.7 to +31.7 seconds. Note, this observed range of performance was from trials conducted under a range of wind conditions and a range of wind forecast errors in the FMS wind entries. The simulation analyses which follow control for these wind factors to explicitly examine their impact.

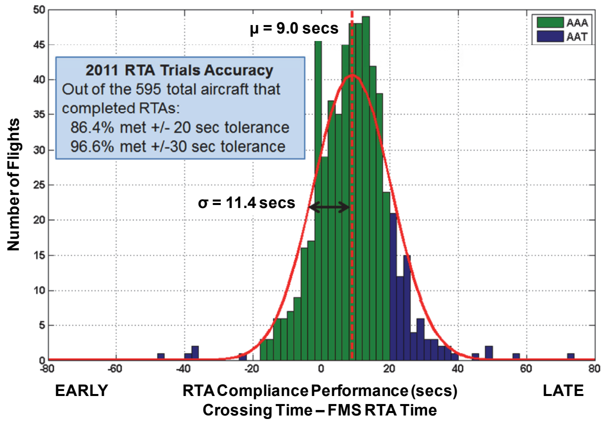

Figure 4. 2011 Seattle 4D-TBO Flight Trials RTA Compliance Performance [7] (AAA = RTA Assigned, Accepted, and Accomplished within programmed 20 secs tolerance, AAT = RTA Assigned, Accepted, and achieved but beyond programmed 20 secs Tolerance)

B. Lincoln/MITRE FMS Simulations

A series of parametric studies were conducted in collaboration between Lincoln and MITRE CAASD using their FMS test benches. Full details can be found in [6], but a summary is included below. Three systems with RTA functionality were available for testing, reflecting a variety of FMS systems in current operational use:

- Boeing B737-700 utilizing a GE FMS (B737/GE) using current operational software version U10.8A in the GE-supplied "sFMS" system.
- Two variants of the B757-200 both utilizing a Honeywell Pegasus FMS in an Aerosim-packaged flight training system adapted for research purposes, including one using current operational ("Black Label") software version PS4083821-910 (B757/HW BL), and one using a "first generation" research prototype ("Red Label v1") software (B757/HW RL) developed to explore adding user-adjustable "RTA tolerance setting", calculation and display of achievable RTA times, and user-adjustable RTA speed limits.

The B757/HW BL variant is representative of an FMS type offering closed-loop RTA speed control in cruise but open-loop RTA speed control in descent, while the other systems have closed-loop RTA speed control in all flight phases.

A range of *Wind Scenarios* were constructed to expose the aircraft to a range of operationally-realistic truth wind conditions, and forecast errors were superimposed to cover 5 and 20 kts Root Mean Square Vector Error (RMSVE). This range is consistent with errors observed from models used in an operational setting. There are a number of ways of quantifying wind information quality, but RMSVE for wind vector (speed and direction) differences or root mean square error (RMSE) for wind speed differences of a forecast relative to truth data were found to be the most commonly used metrics to quantify the performance of operational wind models. For scalar errors, the RMSE is the square root of the average of the individual squared differences between pairs of scalar forecast (*f*) and observation (*o*) quantities, given by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - o_n)^2}$$

where N is the number of forecast-observation pairs. The RMSVE is the RMS error applied to the magnitudes of the forecast and observed wind vector components as given by:

$$RMSVE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(u_f - u_o\right)^2 + \left(v_f - v_o\right)^2}$$

where N is the number of forecast-observation pairs, u is the east-west component of the wind vector, v is the north-south component, and subscripts f and o refer to forecast and observed respectively. In this study, Monte Carlo simulations were conducted (up to 100 runs in each condition). Scenarios were classified as "Benign", "Severe Headwind", and "Severe" based on climatologically representative statistics of wind speed and variability within the local region traversed by the scenario trajectory. Truth wind data for these classifications were obtained from archives of gridded wind numerical weather prediction model data such as the High Resolution Rapid Refresh model (HRRR) and Rapid Refresh Model (RAP). These model analyses provide realistic representations of the winds over the spatial and temporal scales of interest and preserve the spatial correlations between neighboring wind samples. Spatially correlated sequences of simulated forecast errors were created by first generating a normally distributed random error sequence having a mean of zero, and standard deviation (σ) corresponding to the amount of forecast RMSVE being modeled (5 or 20 kts). The initial random error sequence was then filtered with a Gaussian filter kernel having shape parameters consistent with the desired correlation lengths to produce a spatially correlated error sequence (no temporal variation was considered in this analysis).

The *ATC Scenarios* studied all involved flight parameters informed by the 4D-TBO ConOps and the Seattle flight trials, although the geographic location of the simulation runs was not tied to any one location. The ATC scenario trajectories were executed using the test FMSs programmed with appropriate waypoints separated by 10 or 100 NM. For most of the scenarios studied, a straight line lateral path was used, although one case also included turns during descent to expose the aircraft to more rapidly varying truth wind field characteristics. The vertical profiles comprised an initial level cruise segment at FL290 or FL390, followed by a descent to a meter fix at 12,000 ft. An RTA time target was assigned at either 150 NM or 250 NM distance from the fix. The RTA target was set relative to a range of achievable RTA times estimated by the FMS under test (using the forecast winds), as either Mid-range, Early (20 secs after earliest achievable time), or Late (20 secs prior to latest achievable time). Prior to RTA assignment, the aircraft was at a speed governed by the programmed FMS Cost Index, but after RTA assignment, the RTA function of the FMS controlled the aircraft speed as necessary in its effort to arrive at the meter fix to comply with the RTA.

The *Aircraft/Automation Simulations* were conducted using the MITRE FMS simulation capabilities for the B737/GE and B757/HW systems as previously described. The B737/GE and B757/HW RL FMS have user-adjustable RTA tolerance settings (internal FMS sensitivity parameter reflecting time-error value, expressed in seconds, that triggers recalculation of RTA speed targets), while the B757/HW BL FMS has a fixed setting of approximately 30 seconds.

Performance Assessment was measured by comparing the actual time of arrival at the meter fix relative to the target time across multiple runs, from which RTA TE and RTA 95% CI statistics were compiled.

Table I below summarizes the study conditions.

TABLE I. SIMULATION CONDITIONS FOR LINCOLN/MITRE STUDIES

Wind Information Analysis Framework Element	Independent Variable	Values Tested	Number of Permutations
Wind Scenario	Truth field	"Benign" "Severe" with 180° Course Reversal "Severe Headwind"	3
	Forecast error distribution	σ = 5 knots RMSVE σ = 20 knots RMSVE	2
ATC Scenario	Trajectory	Cruise FL290, FL390, Meter fix 12,000 ft	2
	RTA assignment distance (from meter fix	150 NM, 250 NM	2
	RTA time assigned (Relative to range of achievable times)	Early, Mid-range, Late	3
	Waypoint spacing	10 NM, 100 NM	2
Aircraft/Automation Simulation	Aircraft/FMS type	B737/GE (±6 secs RTA tolerance) B757/HW BL (±30 secs RTA tolerance) B757/HW RL (±6 and ±30 secs RTA tolerance)	4
	Total permutations	50-100 Runs/Condition	576

Full details of the analysis can be found in [6], but summary results are provided in Figure 5. A number of observations can be made based on these results, as follows:

- RTA compliance performance was observed to be much better for the B757/HW RL and B737/GE systems compared to the B757/HW BL FMS. The relative performance differences between the systems can be explained by the extent of closed-loop speed control to an RTA during descent. The B757/HW RL and B737/GE systems perform closed-loop speed control in all flight phases while the B757/HW BL only has closed-loop speed control in cruise.
- RTA compliance performance was observed to degrade only slightly with wind forecast error level in the B737/GE and B757/HW RL FMSs, but the B757/HW BL was much more sensitive to wind forecast error.
- The RTA tolerance setting had a medium effect on the RTA compliance performance of the HW RL system which provides for user-adjustable settings, with a slightly bigger effect when wind forecast error was at the 20 kts RMSE level.
- All results demonstrated a tendency to arrive late relative to the RTA on average (no more than 10 seconds for the B737/GE and B757/HW RL systems in this study) which is consistent with results obtained in the Seattle flight trial and other studies of RTA performance. Several hypotheses have been suggested to explain this late bias, but further studies on this point are required to understand and fully characterize this tendency.
- RTA 95% CIs (estimates calculated as $\mu \pm 2\sigma$ (Mean ± 2×Standard Deviation assuming Gaussian distribution) were generally within ±20 seconds of the mean at the tighter 6 sec tolerance setting for all systems with closed-loop speed control in descent under the conditions tested.

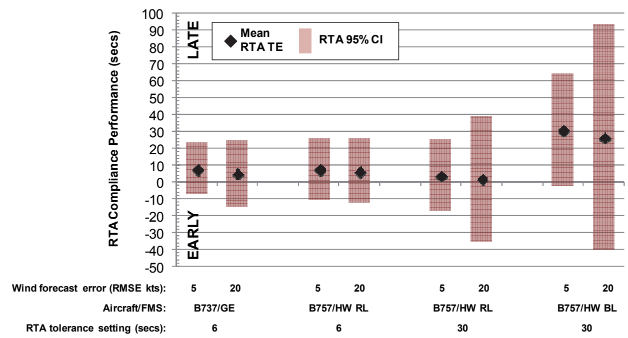

Figure 5. Summary Results from Lincoln/MITRE Simulations

C. *Lincoln FMS Simulation System*

It was desired to supplement the baseline flight trial results and Lincoln/MITRE FMS simulation activities with a flexible simulation system which could be used to model individual components of a 4D-TBO environment, including enhancements to FMS wind-handling capabilities, to better understand their impacts on performance.

An agent-based simulation system was developed at MIT Lincoln Laboratory which provides the capability to establish different levels of fidelity for each element incorporated in the simulation. Agents are created to model individual avionics units, pilots, airline operating centers, or other systems as required to ensure the appropriate characteristics of a particular system are embedded in the simulation to reflect real-world behaviors. The simulation system was also designed to be scalable: each instance can run on a virtual computer allowing multiple simulations to be executed in parallel with additional agents and components able to be started in a cloud infrastructure as needed. The simulation system incorporates both operational and research versions of the Honeywell B757/767 Pegasus FMS. The research version includes software changes undertaken during this research which permit the evaluation of enhancements to wind blending algorithms (WBA) to enable quantification of performance improvement potential for near-term avionics refinements.

The Lincoln FMS Simulation System (LL FMSim) was used to test a range of hypotheses developed to address a range of open questions remaining after the previous simulation study, as follows:

1. Increased magnitude headwind forecast error causes greater magnitude RTA TE.
2. Systems that do not maintain closed-loop control until the RTA fix location have greater magnitude RTA TE as compared to systems that do.
3. Flights at lower cruise levels are less impacted by headwind forecast errors.
4. Headwind forecast errors closer to the RTA fix cause greater magnitude RTA TE.
5. Increased waypoint spacing causes greater magnitude RTA TE.

Full details and results for all the analyses conducted in these five areas can be found in [6], but in the interest of space just the results to assess hypothesis 1 and 4 are included here.

1) Hypothesis 1 Results

To test hypothesis 1, a series of experiments were designed to produce conditions that were expected to simulate varying levels of RTA TE performance. These experiments encompassed simulated flights through realistic wind fields extracted from HRRR data, with wind forecast data at each waypoint determined a priori such that the error in the forecast headwind that the aircraft would experience in its flight would be constant at each waypoint. Wind forecast data were programmed in the FMS at 50 NM intervals, with one cruise-level value for each cruise waypoint, and three for the descent phase. The RTA fix and assigned time were entered 150 NM prior to the fix location. All flights maintained a cruise altitude of FL350 before starting descent to a meter fix at 12,000 ft. Constant headwind forecast errors ranging from +50 kts to -50 kts at 5 kts intervals were tested to simulate an extreme range of forecast error conditions.

Three trials of each experimental condition were conducted to demonstrate the repeatability of the distributed simulation system. To compare the effect if the dominant winds in these trials were either head or tail winds, the experiments were reproduced flying the flight plan in the opposite direction so the experiment would have an equal number of headwind and tailwind cases. All simulations were conducted using the B757/HW RL FMS that maintained closed-loop control until arriving at the RTA fix location.

Results are presented in Figure 6 below. The results indicate that the RTA TE remains relatively low in the tested scenarios (within ± 10 seconds) across the headwind forecast error range of -30 to +15 kts RMSVE. Beyond these thresholds, the magnitude of the RTA TE does increase with increasing headwind forecast error magnitude. From the data, it is clear that overestimating the headwinds on a flight drives the RTA TE towards greater negative values, that is, earlier than desired. Underestimating the headwinds on the same flight causes it to arrive later. Whether the aircraft is flying through a predominantly headwind or tailwind field had little noticeable effect in the RTA TE performance until the headwind forecast exceeded the range stated above.

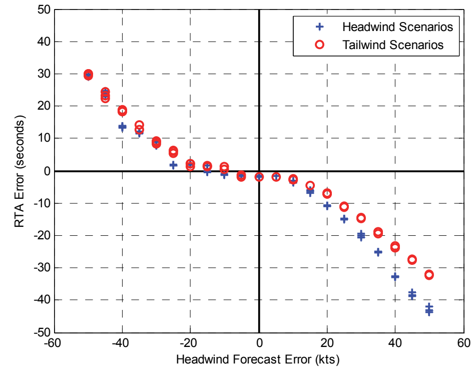

Figure 6. RTA TE as a Function of Constant Wind Error (B757/HW RL FMS)

2) Hypothesis 4 Results

The underlying wind fields used in the experiments were identified by analysis of HRRR data to fall into one of three categories effectively representing wind fields of light, moderate or high wind magnitude conditions. Experiments were conducted with flights through a representatively selected wind profile from each category, but had an additional spatially-correlated wind field variation superimposed on the underlying wind field. The formation of the superimposed wind field was such that it would evaluate into a peak magnitude in wind forecast error that was consistently located at one of four locations in cruise or descent (to vary the peak error distance relative to the meter fix) along the flight plan, as shown in Figure 7.

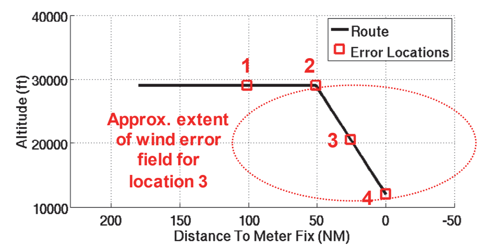

Figure 7. Wind Forecast Error Locations

Fifty randomly generated cases of varying superimposed wind fields were tested for each error location. Three headwind forecast error Gaussian distributions were produced constituting standard deviations with RMSVE values of 5, 10 and 15 kts. Cruise altitude was fixed at FL290 with a meter fix altitude set to 12,000 ft. Waypoint spacing was fixed at 100 NM. Results are presented in Figure 8. The mean RTA TE magnitude does not increase significantly when the peak of the headwind forecast error approaches the RTA fix. However, it was observed that the RTA 95% CI does increase under all error conditions as the peak error location approaches the RTA fix save for the most severe forecast error case tested (RMSVE 15 kts) which had the largest CI span at location 3.

This particularly large CI span is thought to arise because as an aircraft arrives near location 3 in this trajectory, there is insufficient time available for the FMS to correct for a significant forecast wind error. However, further study is required of this effect.

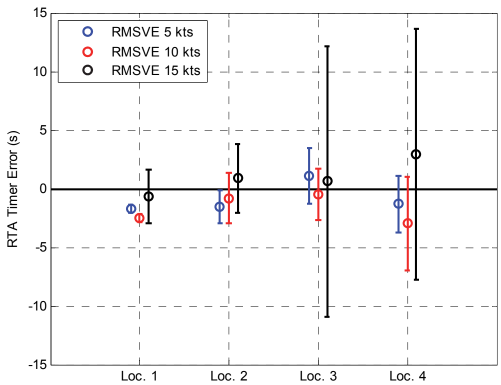

Figure 8. Averaged RTA TE and 95% Confidence Intervals as a Function of Wind Forecast Error Location and Magnitude (B757/HW RL FMS) using 100 NM Waypoint Spacing

D. Key Take-Aways & Associated Performance Trade-Spaces

Based on the results from the analysis scenarios in the previous sections, the relative impact of the key variables on 4D-TBO performance was determined. Each variable was categorized as having a major, medium or minor impact on performance, together with what characteristics of the variable resulted in better or worse performance if considered independently, as shown in Figure 9.

There are some differences between the outputs of the various analyses which warrant further study, extension and validation/verification and that work is on-going. However, the trade-space shown in Figure 10 below shows the range of RTA compliance performance based on a synthesis of the findings from the range of analyses for the specific aircraft/FMS types and scenarios considered to date.

The major performance drivers identified above defined the primary independent variables of the trade-space. The bar heights estimate the possible "best" performance from that combination of variables, while the whiskers reflect the likely range of performance impacts from variations in the other medium and minor performance drivers.

The trade-space reflects the authors' best attempt at synthesizing the findings presented in the preceding sections and is considered to be generally reflective of the relative RTA performance across the range of specific aircraft/FMS, wind and ATC conditions studied. However, care should be exercised in its use for conditions not explicitly studied in this work. Future work will expand the analyses to include more aircraft/FMS types, ATC scenarios and wind scenarios to generate results which cover a broader range of operational conditions.

	Scenario Variable	Overall Impact on Performance	Worse Performance From...*	Better Performance From...*
Aircraft/ Automation	FMS RTA capability	Major	Open-Loop in Descent	Full Closed-Loop Control
	FMS RTA tolerance	Major	Wider RTA tolerance	Tighter RTA tolerance
Wind Scenario	Wind forecast error magnitude	Major	Hi forecast error	Lo forecast error
	Wind forecast error location relative to meter fix	Medium	Near forecast error	Far forecast error
	Truth wind variability	Medium (correlates with error mag)	Hi var truth wind	Lo var truth wind
ATC Scenario	Cruise flight level	Medium	High cruise level	Low cruise level
	Waypoint spacing	Minor	Few cruise wind WPs	Many cruise wind WPs

* All else being equal

Figure 9. Identified Major, Medium & Minor 4D-TBO Performance Drivers

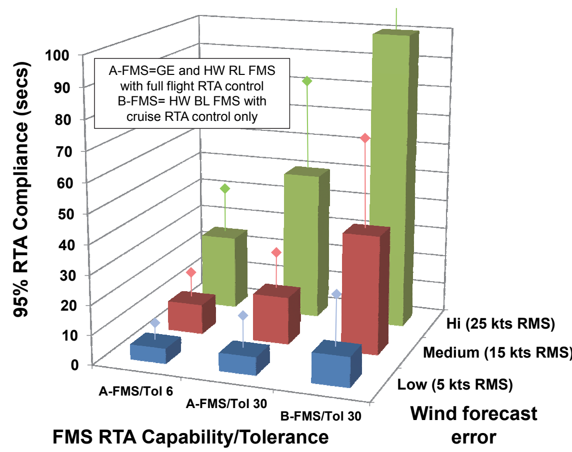

Figure 10. Summary 4D-TBO RTA Compliance Performance Trade-Space for the FMS Types and Scenarios Studied

IV. Analysis of Wind Forecast Model Performance

Results from simulations utilizing the Wind Information Analysis Framework provide the error tolerances that must be satisfied in order to achieve a given target level of performance for the procedure. The forecast error limits will then need to be compared to current and near-term wind forecast model capabilities in order to identify the feasible combinations of forecast model types and their performance as a function of forecast "look-ahead" time (also commonly known as forecast "lead time"), as shown in step 5 of Figure 3.

As documented in prior reports [5,6], a literature survey of publically-available operational wind forecast models was conducted and found outdated, sparse, and inconsistent reporting of model performance with respect to forecast look-ahead time. In order to translate wind forecast error limits to current forecast model capabilities, a more comprehensive, consistent, and updated set of wind forecast model performance statistics was needed. Therefore, an independent analysis of wind forecast model performance was conducted in this work for three operational models used by airline dispatchers and FAA aviation weather systems:

- Global Forecast System (GFS): the GFS model is run by the National Oceanic and Atmospheric Administration's National Centers for Environmental Prediction (NOAA/NCEP) every 6 hours and produces forecast products at two resolutions. For the 0 to 192 hour (8 day) forecast range, the model outputs forecast data on a 25 km horizontal resolution Mercator Cartesian projection with a forecast step resolution of 3 hours and 64 vertical levels.
- Rapid Refresh (RAP): the hourly updating 13 km resolution RAP model replaced the Rapid Update Cycle (RUC) in May 2012 as an operational gridded forecast model produced at NOAA/NCEP. Gridded forecasts of winds and gusts are produced each hour for the North American domain and provide hourly forecast look-ahead steps from 0 to 15 hours at selected altitudes (e.g., 10 meters) and for 50 pressure levels extending to 10 hPa (approximately 100,000 ft under standard atmospheric conditions).
- High Resolution Rapid Refresh (HRRR): the HRRR model is an hourly updating, 3 km resolution, CONUS domain model developed by the NOAA Earth System Research Laboratory (NOAA/ESRL). It became operational at NOAA/NCEP on September 30, 2014. Like the RAP model, the HRRR updates hourly and provides hourly forecast grid sequences of meteorological variables from 0 to 15 hours.

To assess wind forecast model performance, historical GFS, RAP, and HRRR forecast model data were obtained from archives for a 10-month period spanning November, 2013 through August, 2014. In order to represent forecast capabilities across different geographic wind environments, forecast comparisons were made over four separate regions centered on San Francisco (SFO), Phoenix (PHX), Chicago (ORD), and Newark (EWR) airports. Within each approximately 400 NM x 400 NM region, model wind forecasts were sampled and compared against matching wind observations (taken from the matching HRRR 0-hour analysis) at 81 horizontal grid points spaced approximately 50 NM apart, and at nine different pressure altitudes (1000-200 hPa, every 100 hPa, or roughly 350 to 38,600 ft above Mean Sea Level assuming standard atmospheric conditions). RAP and GFS forecasts were laterally interpolated to the finer HRRR grid points. Vertical interpolation was not required as all three models provide wind forecast data grids at the selected pressure levels. Comparisons were made eight times per day and for each of eight selected model forecast look-ahead times (1, 2, 3, 4, 5, 6, 9 and 12 hours). Because the GFS model only has 3-hour forecast look-ahead resolution, comparisons at look-aheads of 1, 2, 4 and 5 hours were made by linearly interpolating in time between the time-bracketing GFS forecasts for those look-ahead times.

Full details of the analysis can be found in [6], but the summary results are shown in Figure 11 below.

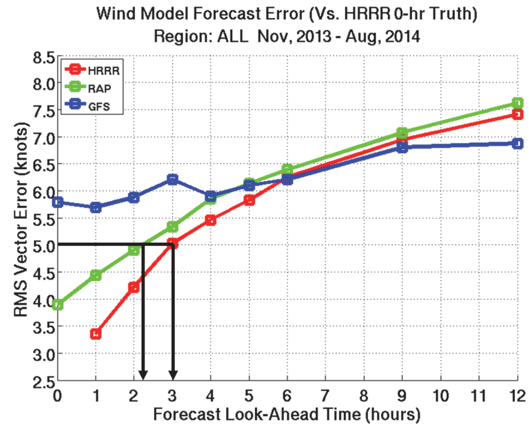

Figure 11. Wind Forecast Model Average Performance Across All Regions

Average wind forecast errors for all three models were found to generally increase with increasing forecast look-ahead time, ranging from 3.4 kts to 7.6 kts over forecast look-aheads of 0 to 12 hours. For forecast look-ahead times of less than 6 hours, the HRRR model produced the best average wind forecast performance, with average RMS vector errors ranging from 3.4 kts at 1 hour look-ahead time to 6.2 kts at 6 hours look-ahead time. The RAP model had average errors ranging from 4.4 kts to 6.4 kts over the same look-ahead time interval. Over look-aheads of 0 to 3 hours, the GFS model has considerably more forecast error (5.2-6.2 kts) than RAP or HRRR. For forecast look-ahead times between 4 and 6 hours, the average forecast errors of the three models generally increases with increasing look-ahead time, but the performance of the three models converges, and by 6 hours, they are comparable (the average GFS performance is slightly better than RAP or HRRR for look-aheads of 6 hours or greater). One possible explanation for the performance similarities at longer forecast look-ahead times is that the numbers and types of upper-air wind observations going into the models are similar across the three models, and at longer forecast look-ahead times, larger and longer scale atmospheric motions and dynamics tend to dominate the numerical forecasts. Differences in treatment of smaller-scale motions and physics along with differences in model grid resolutions come into play more fully for the short-term forecasts. The black lines and arrows in Figure 11 illustrate an example interpretation of the data for determining which model's forecasts would satisfy a hypothetical 5-knot error limit. In the example shown, if a 5-knot error limit is prescribed, then the HRRR forecast look-aheads of up to 3 hours can be expected to provide the required accuracy on average, while only RAP forecast look-aheads of up to 2.1 hours satisfy the requirement. None of the GFS forecasts would meet the hypothetical 5 kts RMSE limit.

Extending beyond the consideration of average forecast performance, Figure 12 shows examples of the forecast error distributions of the three models for the 3, 6, 9, and 12-hour forecasts for all regions analyzed over the full 10 month analysis period. To permit relative comparison, the histogram frequencies were normalized for each model by the maximum frequency of occurrence for that model over the error bins. In the legend, the numbers in parentheses following the model name indicate the total number of forecast comparisons that were performed. Error means and standard deviations are also indicated in the legend.

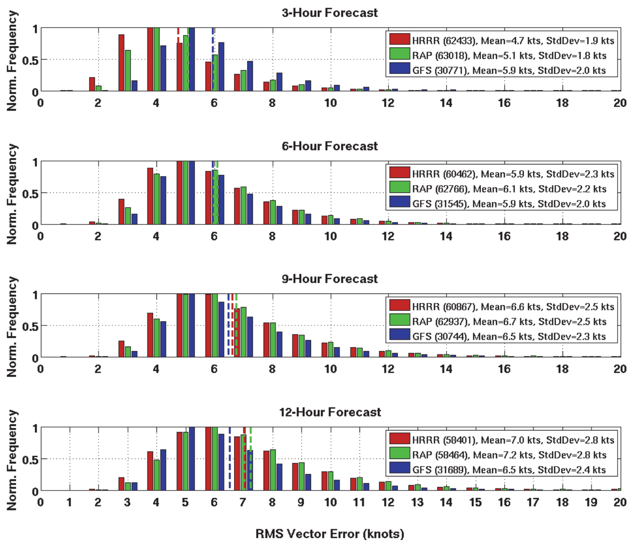

Figure 12. Wind Model Forecast RMS Vector Error Distributions

Considerable spread (standard deviations of 2-3 kts) and long tails are seen in the error distributions, with the tails of the distributions containing errors of 15 kts or more. This implies that forecast errors encountered at any single time or location may be considerably larger than the aggregate means presented in Figure 11, and error limits for a procedure may be exceeded in these instances. These larger wind forecast errors may persist for hours or even days, as seen in Figure 13, which plots a 1-month time sequence of HRRR, RAP, and GFS 3-hour forecast errors from the EWR region at a pressure altitude of 300 hPa (~30,000 ft).

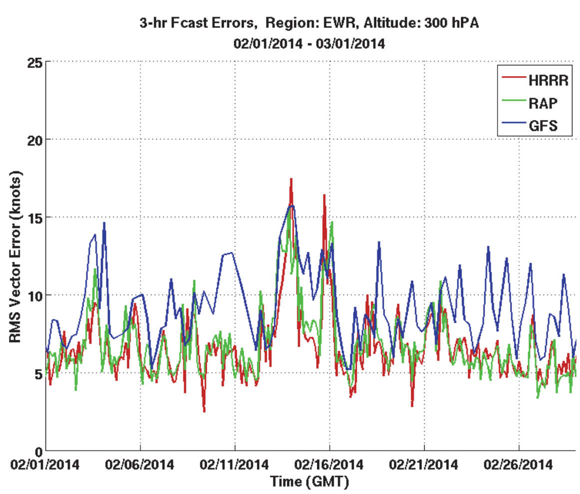

Figure 13. Time Series of 3-Hour Wind Forecast Errors

Note the persistent errors greater than 10 kts in all models during the February 12th-13th time period. Note also that the temporal error trends are very similar across all three models. This is not surprising since the GFS model contributes to the background initialization for the RAP model, and the HRRR model operates as a nested high-resolution grid within the RAP model, and is initialized from within the RAP model. Although the error trends are similar for the three models, there are periods where the GFS exhibits significantly larger errors than the HRRR and RAP models (e.g., around February 11th). These may be periods where the coarser spatio-temporal resolution of the GFS model fails to capture smaller-scale or more rapidly changing wind conditions. More analysis of these time periods is needed to understand the causes of the larger GFS errors.

V. Sample Case Study

In order to demonstrate the utility of the various analyses conducted in the preceding sections, a case study application to establish wind information needs and associated ConOps impacts to support a given level of required 4D-TBO performance is presented. This case study demonstrates the use of the 4D-TBO trade-space from Figure 10 and the average wind forecast model performance results from Figure 11 in the context of the six steps of the Wind Information Implications Flow Diagram. Figure 14 presents the case study. The scenario is a 4D-TBO procedure consistent with the analyses reported earlier. This allowed the trade-space presented in Figure 10 to be used for this case. From this trade-space, combinations of FMS capability and wind forecast error which achieved a ±10 secs (i.e., 20 secs 95% CI assuming zero mean) performance were identified. This target was chosen to reflect current draft performance standards being considered by the community, which our results suggest could be a challenge to achieve under certain scenarios. Assuming a desire to enable a procedure which could be supported by any FMS capability, a need for wind forecast error <5 kts RMSVE was identified from the trade-space. Then referring to the wind forecast model average performance as a function of lookahead time summarized in Figure 11, a need for the FMS to be using HRRR data less than 3 hours old, or RAP data less than 2.1 hours old on average was identified (the GFS model cannot support this error level at any lookahead time on average). These findings imply a need to deliver wind data to an aircraft that was less than these age requirements. For a short-haul flight of less than 2 hours with RAP data, or less than 3 hours for HRRR data, pre-flight winds loaded shortly before departure could support a 4D-TBO operation at the ±10 secs 95% of time RTA compliance performance level on average (performance on any given day may differ due to the range of model performance highlighted in Figures 12 and 13). Flights with longer durations would require wind uplinks en route to support this level of performance (possibly multiple wind uplinks for long-haul flights if the aircraft was required to provide a valid ETA window prior to a ground system establishing a feasible meter fix target time).

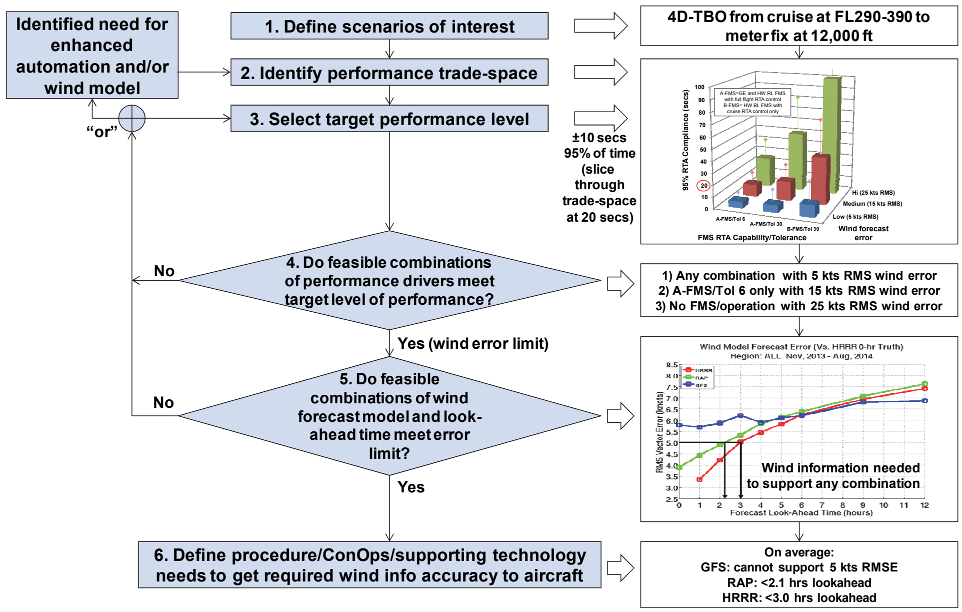

Figure 14. 4D-TBO Example Case Study

VI. Summary

This paper summarizes work designed to: (1) Establish the relationship of wind information accuracy to 4D-TBO performance for a selection of operationally-relevant scenarios to identify wind needs to support them, and (2) Present examples of how this information can be used to determine what wind information content and update rate to the aircraft will deliver a given target performance level to help inform concept of operations development and supporting technology needs.

The research was informed by significant stakeholder input to determine the scenarios and metrics of highest potential value. The outputs of the study have also been regularly briefed to a range of interested RTCA committees (e.g. those defining performance standards and ConOps in relevant areas) for feedback and incorporation into their work areas if appropriate. Fruitful collaborations and technical interchanges have also occurred with FMS and airframe manufacturers which have informed this work and identified potential high value avionics and wind information improvements. Through these avenues, this work is helping advance the dialog of wind information needs to support NextGen applications.

Acknowledgments

This work was sponsored by the Federal Aviation Administration (FAA) under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government. Many thanks to Gary Pokodner, Eldridge Frazier, Steve Abelman and Rick Heuwinkel in the Weather Technology in the Cockpit group and Aviation Weather Division for their guidance and support. Thanks to Will Symionow and Tom Becher at MITRE CAASD for support with FMS simulations, as well as Mike Jackson and Ryan Howe-Veenstra at Honeywell for FMS collaborations. Thanks to Lanie Sandberg, Chris Edwards, Tan Trinh and Yan Glina at MIT Lincoln Laboratory for their work on the project team.

References

[1] De Smedt, D. & G. Berz, "Study of the Required Time of Arrival Function of Current FMS in an ATM Context", *Digital Avionics Systems Conference,* 2007.

[2] Klooster, J. K., A. Del Amo & P. Manzi, "Controlled Time-of-Arrival Flight Trials: Results and Analysis", *8th USA/Europe Air Traffic Management Research and Development Seminar* (ATM2009), Napa, CA.

[3] Balakrishna, M., T. A. Becher, P. V. MacWilliams, J. K. Klooster, W. D. Kuiper & P. J. Smith, "Seattle Required Time-Of-Arrival Flight Trials", *Digital Avionics Systems Conference* (DASC), Seattle, WA, 2011.

[4] Vaddi, S., P. Sengupta, M. Tandale & J. E. Robinson, "Large-Scale Data Analysis for Characterization of the Effect of Wind Forecast Errors on ETAs", *14th AIAA Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, 2014.

[5] Glina, Y., T. G. Reynolds, S. Troxel & M. McPartland, "Wind Information Requirements to Support Four Dimensional Trajectory-Based Operations", *12th AIAA Aviation Technology, Integration, and Operations Conference*, Indianapolis, IN, AIAA 2012-5702, 2012.

[6] Edwards, C., M. D. McPartland, T. G. Reynolds, M. J. Sandberg, T. L. Teller & S. W. Troxel, "Wind Information Requirements for NextGen Applications: Phase 3 Final Report", MIT Lincoln Laboratory Report ATC-422, 2014, http://www.ll.mit.edu/mission/aviation/publications/publication-files/atc-reports/Edwards-C_2014-ATC-422_RR-104746.pdf.

[7] Wynnyk, C., and D. Gouldey, "2011 Seattle Required Time of Arrival (RTA) Flight Trials Analysis Report," MITRE CAASD, July 2012.

Author Biographies

Tom Reynolds is Assistant Leader of the Air Traffic Control Systems Group at MIT Lincoln Laboratory. He has a PhD in Aerospace Systems from MIT.

Michael McPartland is Technical Staff in the Air Traffic Control Systems Group at MIT Lincoln Laboratory. He has a PhD in Aerospace Engineering from the State University of New York at Buffalo.

Tom Teller is Associate Staff in the Surveillance Systems Group at MIT Lincoln Laboratory. He has a BS in Aerospace Engineering from Penn State.

Seth Troxel is a sub-contractor with the Air Traffic Control Systems Group at MIT Lincoln Laboratory. He has a BS in Meteorology from San Jose State University.