Project Report
ATC-431

Wind Information Requirements for NextGen Applications Phase 4 Final Report

C. Edwards
Y. Glina
J.C. Jones
M.D. McPartland
T.G. Reynolds
S.W. Troxel

6 March 2017

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Federal Aviation Administration,
Washington, DC 20591

This document is available to the public through
the National Technical Information Service,
Springfield, Virginia 22161

This material is based upon work supported by the Federal Aviation Administration (FAA) under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the FAA.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
ATC-431		

4. Title and Subtitle	5. Report Date
Wind Information Requirements for NextGen Applications, Phase 4 Final Report	6 March 2017
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
C. Edwards, Y. Glina, J. Jones, M. McPartland, T. Reynolds, and S. Troxel	ATC-431

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
MIT Lincoln Laboratory 244 Wood Street Lexington, MA 02421	11. Contract or Grant No. FA8721-05-C-0002 and/or FA8702-15-D-0001

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Department of Transportation Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	Project Report
	14. Sponsoring Agency Code

15. Supplementary Notes

This report is based on studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology, under Air Force Contract FA8721-05-C-0002 and/or FA8702-15-D-0001.

16. Abstract

The success of many NextGen applications with time-based control elements, such as Required Time of Arrival (RTA) at a meter fix under 4D-Trajectory Based Operations (4D-TBO)/Time of Arrival Control (TOAC) procedures or compliance to an Assigned Spacing Goal (ASG) between aircraft under Interval Management (IM) procedures, are subject to the quality of the atmospheric forecast utilized by participating aircraft.

Erroneous information derived from provided forecast data, such as the magnitude of future headwinds relative to the headwinds actually experienced during flight, or forecast data that is insufficient to fully describe the forthcoming atmospheric conditions, can significantly degrade the performance of an attempted procedure.

The work described in this report summarizes the major activities conducted in Fiscal Year 2015.

17. Key Words	18. Distribution Statement
	Approved for public release: distribution unlimited. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	106	

FORM DOT F 1700.7 (8-72) Reproduction of completed page authorized

This page intentionally left blank.

REVISION HISTORY

Version	Description	Release Date
0.0-0.x	Internal development versions	N/A
1.0		
2.0		

This page intentionally left blank.

ACKNOWLEDGMENTS

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.

The authors extend their appreciation to Airlines for America and Greg Pratt of the National Oceanic and Atmospheric Administration (NOAA) Global Systems Division for their assistance with the Meteorological Data Collection and Reporting System (MDCRS) aircraft meteorological data used in this study, Chief Technical Pilot Rocky Stone and Chris Osterman of United Airlines for their assistance in providing simulated flight data, and Captain Alan Midkiff of American Airlines for flight data and technical advice.

This page intentionally left blank.

EXECUTIVE SUMMARY

The success of many NextGen applications with time-based control elements, such as Required Time of Arrival (RTA) at a meter fix under 4D-Trajectory Based Operations (4D-TBO)/Time of Arrival Control (TOAC) procedures or compliance to an Assigned Spacing Goal (ASG) between aircraft under Interval Management (IM) procedures, are subject to the quality of the atmospheric forecast utilized by participating aircraft.

Erroneous information derived from provided forecast data, such as the magnitude of future headwinds relative to the headwinds actually experienced during flight, or forecast data that is insufficient to fully describe the forthcoming atmospheric conditions, can significantly degrade the performance of an attempted procedure.

The work described in this report summarizes the major activities conducted in Fiscal Year 2015 (FY15), which builds upon prior work. The major objectives were:

1. Enhance the existing Wind Information Analysis Framework (WIAF) to
    a. Support increased numbers of simultaneous aircraft simulations
        i. Major system-level architectural changes were made to double the number of simulations that could be run in parallel from 20 to 40.
    b. Augment agent models
        i. Many updates were made to the agents; for example, in the pilot agent, we added speed error handling, reacting to Flight Management System (FMS) messages, spoiler control modeling to simulate actual inflight usage, etc.
    c. Create the ability to simulate actual flights and apply actual atmospheric conditions
        i. Fly simulations of actual flights, utilizing weather data from the aircraft itself to reproduce realistic experienced weather conditions.
2. Analyze the performance of publically available forecast as compared to in situ reported atmospheric conditions
    a. Comparisons of High Resolution Rapid Refresh Model (HRRR) forecast data to recorded in-flight weather data from the Meteorological Data Collection and Reporting System (MDCRS).

3. Support Radio Technical Commission for Aeronautics (RTCA) Special Committee needs, especially for SC-206, "Aeronautical Information and Meteorological Data Link Services," including co-leading Sub-Group 7 created to develop "Guidance for Data Linking Forecast and Real-Time Wind Information to Aircraft."

4. Determine if augmented FMS wind-handling capabilities provide a meaningful improvement in RTA performance tested under real-world Air Traffic Control (ATC) and atmospheric conditions

    a. Simulations were run in three conditions: (1) no forecast data, (2) forecast data at four levels of descent, and (3) forecast data at nine levels of descent.

5. Provide recommendations for high value future work.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS
(Continued)

LIST OF ILLUSTRATIONS
(Continued)

LIST OF ILLUSTRATIONS
(Continued)

LIST OF TABLES

This page intentionally left blank.

1. INTRODUCTION

1.1 MOTIVATION

Many NextGen applications depend on access to accurate forecasted wind data. Applications such as Required Time of Arrival (RTA) at a meter fix under 4D-Trajectory Based Operations (4D-TBO)/Time of Arrival Control (TOAC) procedures or compliance to an Assigned Spacing Goal (ASG) between aircraft under Interval Management (IM) procedures each rely on accurate representation of flown winds. Figure 1 illustrates how wind information is used by Air Traffic Control (ATC) on the ground to develop time targets for use in a 4D-TBO procedure. Wind information in the aircraft is used by the Flight Management System (FMS) or other 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. Note that the mean error may be zero or slightly offset. Target performance is likely to be specified as a maximum allowable RTA compliance error expected for a given fraction of operations, for example ±*x* seconds 95% of the time [1]. Any errors in the aircraft wind information relative to the truth winds actually flown through can potentially degrade the performance of the procedure. 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 increasing the resolution of the forecast model in the relevant avionics system.

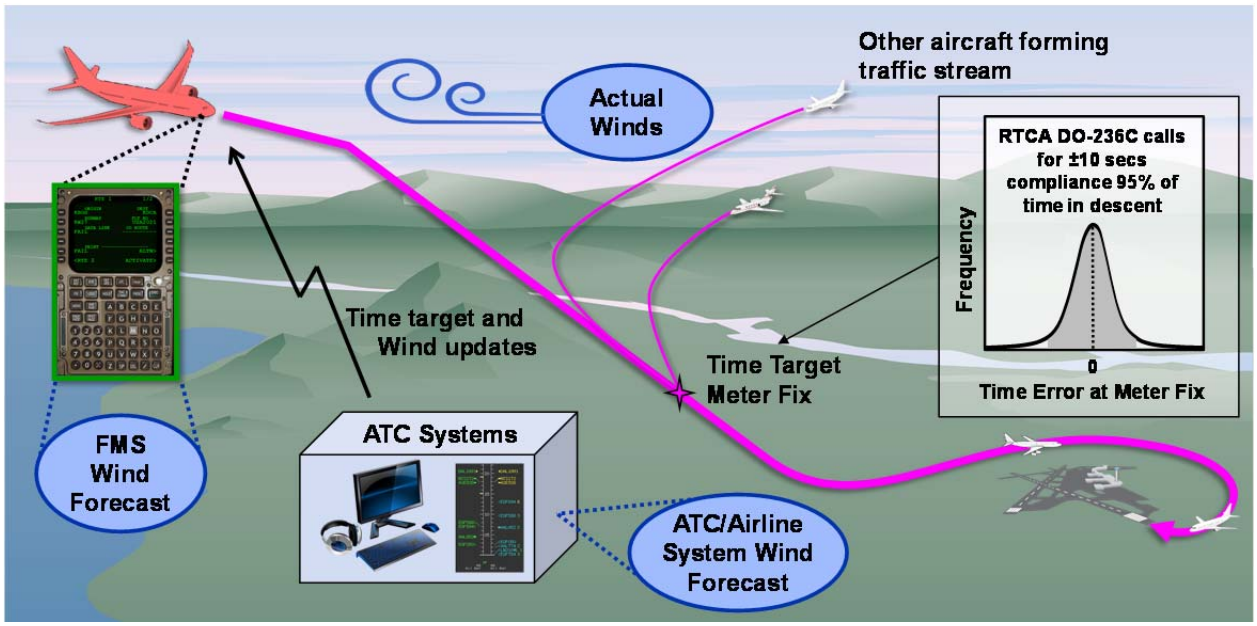

Figure 1. Focal elements relevant to 4D-TBO operations.

1.2 SUMMARY OF PRIOR WORK

In Phase 1 of this work (corresponding to FY12), a generic Wind Information Analysis Framework was developed to explore wind information needs across a range of NextGen applications. The framework was applied to a 4D-TBO scenario to act as a "proof of concept" of its use. It illustrated that even simplified executions of its elements could yield interesting and complex results, which could be of high value in determining how 4D-TBO performance varies with wind information quality. Phase 2 of the work (largely corresponding to FY13) built upon this foundation by using refined and expanded applications of the Wind Information Analysis Framework. It included tasks to (1) increase modeling fidelity and explore more complex 4D-TBO procedures; (2) expand the set of wind forecast scenarios and metrics; (3) assess performance of 4D-TBO with realistic future FMS wind-handling enhancements; and (4) expand the focus applications to include Interval Management (IM), both Ground-based Interval Management (GIM) and Flight-deck Interval Management (FIM). Prior work has also undertaken extensive assessment of wind information quality metrics, as well as the performance of a range of wind forecast models used by aviation stakeholders in the U.S. and overseas. Principal outcomes from Phase 3 of this work included (1) analysis of the impact of wind information on 4D-TBO and IM performance of synthetic routes in synthetic environments, (2) analysis of various publically available wind information products available for use in the wind implications process flow diagram, and (3) example case studies of implications of different wind forecast error limits on 4D-TBO and IM trade-spaces. Full details of all this work can be found in [2–6].

1.3 CURRENT RESEARCH ACTIVITIES AND DOCUMENT OUTLINE

The Phase 4 work summarized in this report has the objective of building upon the Wind Information Analysis Framework with application to focus areas from earlier phases to help establish wind information needs for a range of NextGen applications to directly support priorities of the sponsor and stakeholder community. The sections of the report are organized as follows:

- **Section 2** summarizes augmentations to the analysis infrastructure, including refinements and extensions to the Wind Information Analysis Framework (WIAF) as well as the creation of the Meteorological and Flight Information Database (MAFID), which integrates with the WIAF, resulting in a highly flexible and scalable simulation and analysis infrastructure to support the current and future efforts.

- **Section 3** presents further analysis of select publically available forecast models and aircraft-derived atmospheric measurements, which build upon prior work and tailor the analysis for the current objectives.

- **Section 4** describes the role and activities undertaken in this phase of the program in support of RTCA Special Committees' needs.

- **Section 5** presents analysis of the effect of augmentations to the Boeing 757/767 Pegasus FMS (in terms of additional descent forecast wind levels) designed to improve RTA performance, which addresses sponsor and RTCA needs by leveraging the analysis infrastructure developed.

- **Section 6** presents a summary of the report and recommends next steps to refine and extend this work.

Note this work is not intended to specifically *recommend* concepts of operation and/or datalink technologies to support 4D-TBO or IM applications, but rather to identify the wind sensitivities of these applications and provide a process by which this information can be used by stakeholders to assess implications for operation and/or datalink technologies.

This page intentionally left blank.

2. ENHANCED ANALYSIS INFRASTRUCTURES

2.1 INTRODUCTION

In order to help explore the relationship of wind information to NextGen application performance, a Wind Information Analysis Framework has been developed and refined throughout the various phases of this work. The latest version of the framework is shown in Figure 2.

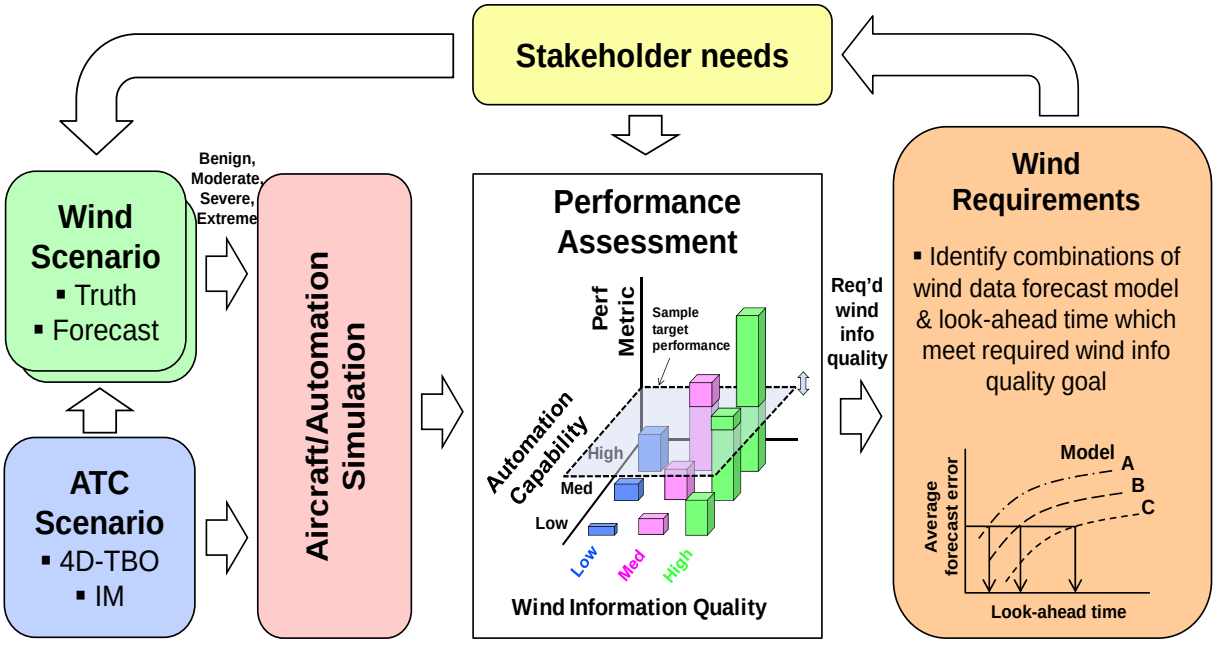

Figure 2. Wind Information Analysis Framework.

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 and pressure. The **Aircraft/Automation Simulation** element represents the behavior of the aircraft, engine, autopilot, and Flight Management System (FMS) in the context of the wind scenario and ATC application being studied. By running simulations of how aircraft perform in the context of a given ATC application when given varying qualities of wind forecasts when 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 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 aircraft 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 capability that exceed that standard can be identified. The **Wind Requirements** element identifies which combinations of wind data content, from which specific operational wind forecast models at what forecast 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 procedure performance. 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 to date were a range of RTCA Special Committees (SC-206, 214, 227, and 186) with representation across the FAA, airlines, and industry. Future work will continue to broaden stakeholder engagement as appropriate for the work.

2.2 IMPROVEMENTS TO INFRASTRUCTURE

2.2.1 Meteorological and Flight Information Database (MAFID)

A number of new capabilities were required to support the Phase 4 research. These new capabilities would require the development of means to collect, store, investigate, and disseminate meteorological and flight data relevant to the analyses. This set of systems and services developed for this program are colloquially known as Meteorological and Flight Information Database (MAFID).

Identification of Data Sources

The principal data sources leveraged in this research can be categorized into the following types:

- Meteorological measurements
  - Meteorological Data Collection and Reporting System (MDCRS) – Collects and stores real-time automated position and weather reports from participating aircraft.
  - Radiosondes – A battery-powered telemetry instrument package carried into the atmosphere usually by a weather balloon that measures various atmospheric parameters and transmits them by radio to a ground receiver.
- Meteorological forecasts

  - Publically available numerical weather prediction models including the Global Forecast System (GFS) and High Resolution Rapid Refresh (HRRR).

- Flight planning information
  - Traffic Flow Management System (TFMS) – Data stream service for supporting the management and monitoring of national air traffic flow and includes reports on flight plan filing, amendments, cancelations, etc.

- Flight surveillance information
  - Airport Surface Detection Equipment, Model X (ASDE-X) – Used for airport terminal and surface area detections and contains aircraft information from aircraft transponders used by air traffic controllers to identify aircraft.
  - Traffic Flow Management System (TFMS) track – a part of the data stream service with aircraft positional reports.

- Navigational data
  - Coded instrument flight procedures (CIFP) – An FAA-provided navigational database used as a basis to support area navigation.

These data types, whether static or streaming, were brought into MAFID by developing and deploying the appropriate software and hardware infrastructure to permit their incorporation into the underlying database. The architecture of the database and the web-based MAFID services that leverage it were developed to support multiple sources for similar data types (e.g., multiple meteorological forecast models).

As part of the MAFID system, there exists a collection of software applications deployed for continuous streaming and archiving for a number of the data feeds. Additional applications are deployed on independent systems to monitor data streams and system components to ensure a high level of quality of service. Consumed data are processed by a collection of data adapters that interpret, filter, and value add with calculated metadata before ingesting into the underlying database.

MAFID Services

Deployed on top of the underlying database are a set of services. The services are used for analysis and scenario creation.

Route Navigational Service

To support this investigation, a web-based routing navigational service was developed that performs functions normally found in ground-based ATC systems and internally to FMS systems. This service, amongst other things, provides the means to decode En Route Automation Modernization

(ERAM) generated flight plans and provide detailed characteristics of waypoints, fixes, routes, and airport information required by other subsystems and processes in this study.

Weather Forecast Data Service

One of the principal elements of MAFID is providing a web-based weather forecast data service. An interesting aspect of this forecast data service is that it also operates as a historical forecast data service. That is, the system may be queried for forecast information that would have been available to a request that was made in the past. Another feature of this service is that it is designed as a navigational forecast data service optimized for providing weather data at points along a track projected across time and space.

A more detailed description of the Weather Forecast Data Service and the underlying data infrastructure that supports it is given later.

2.2.2 Improvements to the WIAF

As an outcome of recommended work from earlier work and as determined during this current effort and with particular respect to RTCA support, a number of infrastructure enhancements were necessary. The incorporation of these enhancements led to significant improvements in terms of capabilities, quality, and expediency of analysis.

Additional Number of Aircraft Systems

The agent-based simulation system developed at Lincoln Laboratory under previous phases of work provides the capability to establish different levels of fidelity for each element incorporated in the simulation. The simulation system incorporates both operational and research versions of the Honeywell B757/767 Pegasus FMS. 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 with the concept of scalability from the outset.

There is a constraint with the provided Honeywell FMS in that, since it operates as a Windows service, only one instance can execute on a given Windows PC at one time. As such, each FMS instance is run on a virtual computer inside a Lincoln Laboratory cloud infrastructure allowing multiple simulations to be executed in parallel along with its required agents and components to enable quasi-fast-time results generation.

In this phase of work, a contracted license change with Honeywell permitted the doubling of the number of simultaneous FMS instances to run in the framework from 20 to 40.

As part of the framework augmentation, all the agents associated with a single aircraft instance (GPS, Multifunction Control Display Unit [MCDU], Attitude Heading Reference System [AHRS], etc.) were moved from operating on an associated physical or virtual Linux instance to the same virtual

Windows instance that hosts an aircraft's Honeywell FMS. This organizational change has made scaling the infrastructure straightforward and also simplified system management.

Improved System Models

Engines

It has been identified that the default Rolls-Royce engine model used in the WIAF was accurate in terms of fuel consumption and thrust generation modeling during cruise operations, but the accuracy degrades at lower altitudes and in particular at low throttle conditions. The FMS maintains its own engine model for the particular aircraft in which it operates. If the characteristics of the engine model in the FMS differ significantly from the actual/simulated engine, there would likely be degradation in overall RTA performance.

In order to improve the engine model, the original model, which was provided in compiled form from the vendor, had to be reengineered. The software model was disassembled into assembly code and the parameters used to define the model identified.

To remodel the engine, flight data for the same model airframe and engine model used in the simulation was obtained. These data included recorded engine and flight conditions at high cruise, mid-altitude, and descent to surface conditions.

A representation of the assembled binary was modelled in MATLAB, which provided the ability to reproduce the engine behaviors as a function of the 290 distinctive coefficients that are used to define the engine model. Having actual flight data as truth data enabled a comparison between the engine model's performance to the flight conditions, which in turn permitted refinements to the model coefficients to match the flight data. A nonlinear optimization with both linear and nonlinear inequality constraints was formulated and applied against the model coefficients to minimize the performance difference observed through the various domains of flight. Fortunately, through analysis, the set of coefficients necessary to consider in the optimization was reduced to 147 design variables of which only a subset were adjusted to improve the model's performance. After the optimization was complete, the new design coefficients were converted to assembly code and reassembled into an executable engine model.

The thrust performance is tightly coupled with the fuel flow in a modern turbofan engine, so a fair way to compare the performance of an engine model with an actual system is to look at the fuel flow under various conditions of flight. Table 1 shows the comparison of some idle-throttle fuel flow data taken from flight data and recalculated using the original and optimized engine models. The idle-throttle conditions are demonstrated because this is where we have observed the largest error in the original model. The optimized model is seen to have significantly lower errors at this condition than the original model.

TABLE 1

Original and Optimized Engine Model Performance Errors

Pressure Altitude (ft)	Mach No.	Fuel Flow (PPH)	Original Model Error (PPH)	Optimized Model Error (PPH)
35000	.780	900	371	0.04
25000	.650	1000	527	–0.18
9632	.484	1023	380	–0.05
6560	.422	1120	387	0.14
3584	.337	1239	478	–3.30
1536	.268	1288	609	8.40
0	.131	1239	392	–59.3
0	.122	1355	487	53.5
0	0	1355	317	0.84

Pilot Agent

A need was identified during the previous phase of work for a reactive pilot agent. This modified agent would need to apply a heuristic to emulate several actions an actual pilot might conduct during a typical flight. The most important changes relative to this work include the agent's ability to react to speed errors during descent, react to FMS DRAG REQUIRED messages, and dynamically set the altitude correction ("baro" setting) to the field station pressure when transitioning out of Class A airspace. Additional behaviors implemented included the setting of appropriate flap positions and lowering gear during the appropriate stage and condition of flight.

The circumstances of when and how the pilot agent responds to the various conditions were developed with the aid of an active B757 captain with over 20,000 hours of flight-time experience. In brief, the pilot agent would deploy spoilers up to 75% to slow the aircraft during idle or near idle descents if observing that the target airspeed was greater than a certain amount for a certain period of time. Spoilers might be raised, lowered, or stowed depending on the response of the system. The reactiveness in terms of delay, thresholds, release hysteresis, and other related parameters varied as a function of flight state. Typically, the reactivity of the pilot agent increased as altitude decreased.

Responding to DRAG REQUIRED messages from the FMS was also implemented. A heuristic was also created to account for this FMS recommendation and is similar though does not equate to the behaviors for the use of spoilers as a function of speed errors. The DRAG REQUIRED message may be presented under various circumstances by the FMS and not just to reduce speed. It may be presented in order to maintain a certain descent profile even if the current indicated airspeed is lower than the target airspeed and the throttles are not at idle.

The application of the pilot agent to set the destination station pressure in the Kollsman window at the appropriate stage of flight captures a condition that is often overlooked in many simulations. It is particularly important in Vertical Navigation (VNAV) descents because, if the station pressure is anything other than 29.92 in-Hg, a discontinuity of the vertical descent profile is created. This discontinuity may be hundreds of feet and the aircraft will have to react accordingly, and as such, there is an effect on the overall performance of the system.

Autothrottle

One area of known limitations is the closed-loop performance of the existing autothrottle model. In the WIAF, the Thrust Management Computer (TMC) is separated into two parts. One part is the digital state machine and the second is the closed-loop speed controller. The TMC state machine was developed as part of this program, but the closed-loop speed controller was left as the original that was delivered with the Lockheed Martin aerodynamically modelling system. Evaluation of the speed controller's dynamics indicate that the existing control laws sufficiently model the actual TMC performance at high altitudes but respond considerably slower than the actual system at lower altitudes.

Figure 3 shows the response to step changes in commanded speed at Flight Level (FL) 390 and at 15,000 feet for both an actual B757-200 and the simulated aircraft under the same weather conditions. In subplots A and B at FL390, the frequency response of the 15 knot speed changes of the simulated system reasonably follows the responses observed in the flight data. However, in subplots C and D, the frequency response to the 20 knots speed changes is significantly lower.

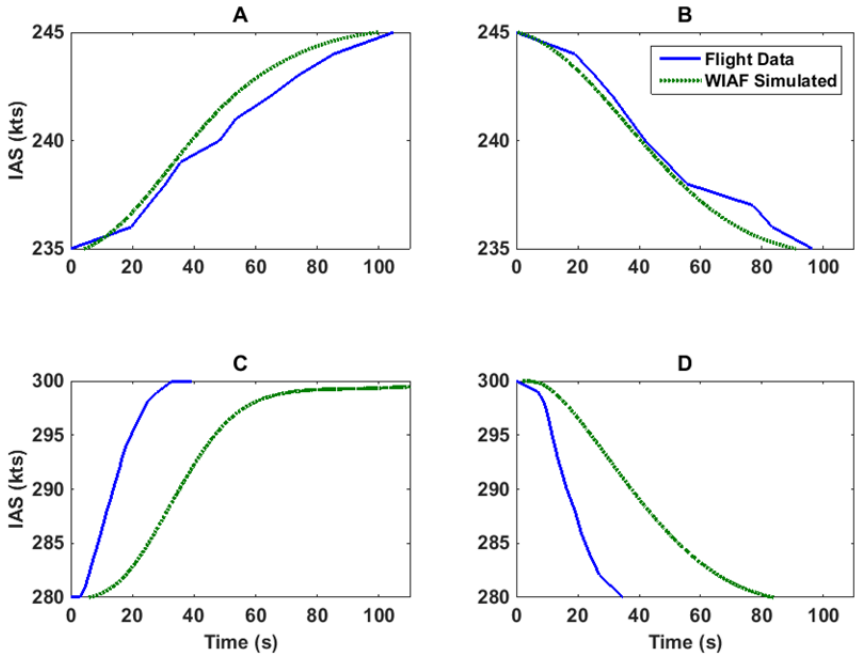

Figure 3. Step responses to command speed changes at FL390 (plots A and B) and 15,000 ft MSL (plots C and D).

To remedy this performance difference, the closed-loop control laws for the TMC need to be remodeled. Sufficient data to do so has been collected from flight data and from a United Airlines FAA Level-D training system that underwent a set of controlled tests to scenarios to excite the principal dynamics of the Thales TMC system utilized in that make aircraft. Due to other priorities, the new control system has not been developed during this phase of effort and is left for future work.

The subpar performance of the simulated control system will have a negative impact on the RTA performance of the system under test. The degree of this impact has not yet been evaluated.

Enhancements to FMS Capabilities

The FMS supplier Honeywell had already provided to Lincoln Laboratory (LL) for incorporation into the WIAF the current operational ("Black Label") software version PS4083821-910 of the Pegasus FMS for the Boeing B757-200. A second version of this software (a "Red Label" candidate in certification parlance), which included closed-loop speed control for RTA operation through all phases of flight, was also provided. This second version was modified for LL during the previous phase of this project as a research variant to employ two proprietary wind-blending algorithms, which were selectable and analyzed as part of previous work [6].

A particular question of interest raised from previous work was "Do more levels of forecast descent winds improve RTA performance?" To answer that question, Honeywell was contracted under this phase of work to augment the flight code in the latest LL research variant of the Pegasus FMS to increase the maximum number of entered descent wind levels from four to nine. An analysis of the effect the number of levels on performance will be is presented later in this document.

An additional enhancement implemented in the modified FMS was the ability to explicitly specify the forecasted temperatures at each descent forecast altitude. This capability, though utilized as part of this work, has not been evaluated at this time.

2.2.3 High-Fidelity Atmospheric Modelling

In previous work, the atmosphere through which simulated aircraft flew was derived from HRRR zero-hour analysis data (see Section 3.2.2). This forecast model was clearly the best choice at the time to use in our earlier simulations as a representation of the atmosphere. It provided a 3 km gridded data surface at 50 pressure levels over the contiguous United States (CONUS). The zero-hour analysis is based on assimilation of the most recently acquired sensor measurements (including balloon soundings and aircraft reports) together with an initial estimate based on the prior forecast, and represents the model's best estimate of the current state of the atmosphere at that specific time.

It is understood that the HRRR model, like its predecessors, is limited to express vertical and horizontal wind shears (e.g., rapidly changing wind shear across altitudes) due to the generalized mesoscale modeling employed and to some degree by limitations of the dimensional step sizes that define the spatial frequencies that can be expressed in the data.

Correct expressions of conditions that change over time are also limited when using hourly forecast models. One of the problems is that even if two adjacent zero-hour, or future-hour, forecasts accurately model the conditions at the forecast times, the weather conditions occurring between the two forecast times are indeterminate.

Consider the case where the forecast winds at a point in space are expected to be 30 knots from the northwest at some forecast time and 30 knots from the east one hour later. Such a condition could occur when a strong, rapidly moving cold front passes across the region. Unfortunately, there is insufficient information in an hourly model that tells when the front passes the point of interest. Is it 10 minutes past the first forecast time, or 50 minutes past the first forecast time?

Without higher-order derivative information of the states in the forecast, there is little recourse on what one can assume on how the system is changing between the forecast times. Thus, in most cases, one assumes a linear relationship and applies a linear interpolation between the two forecasts to estimate a condition.

This technique has been applied in our previous work and is still utilized in our weather forecasting service (see below) for estimating conditions between forecast times.

Because of these limitations, a new approach was adopted to estimating atmospheric conditions applied to the aircraft during simulations in order to best emulate the dynamic wind conditions that are experienced during flight. Additionally, the approach is correlated to actual date and times so forecasting data available at those times from different systems can be evaluated.

Aircraft Reported Atmospheric Model (ARAM)

To maximize the fidelity of the simulations and the test conditions, aircraft-derived meteorological data reports from MDCRS were used. These data are used in concert with the actual flight data, (e.g., flight plan, cruise altitudes and speed, etc.) that correlates to the particular flight of interest.

In these experiments, the simulated aircraft are placed at correlated positions of space and time that match to an instance of the reported flight data and are thereafter applied the same atmospheric conditions during simulation as observed during the original flight as it spatially transgress the same route as defined by their flight plan.

There is significant effort required in order to collect and identify aircraft reported atmospheric measurements with flight information that describes the route of flight that was assigned to the reporting aircraft. The process to do so is outlined in the individual steps shown in Figure 4.

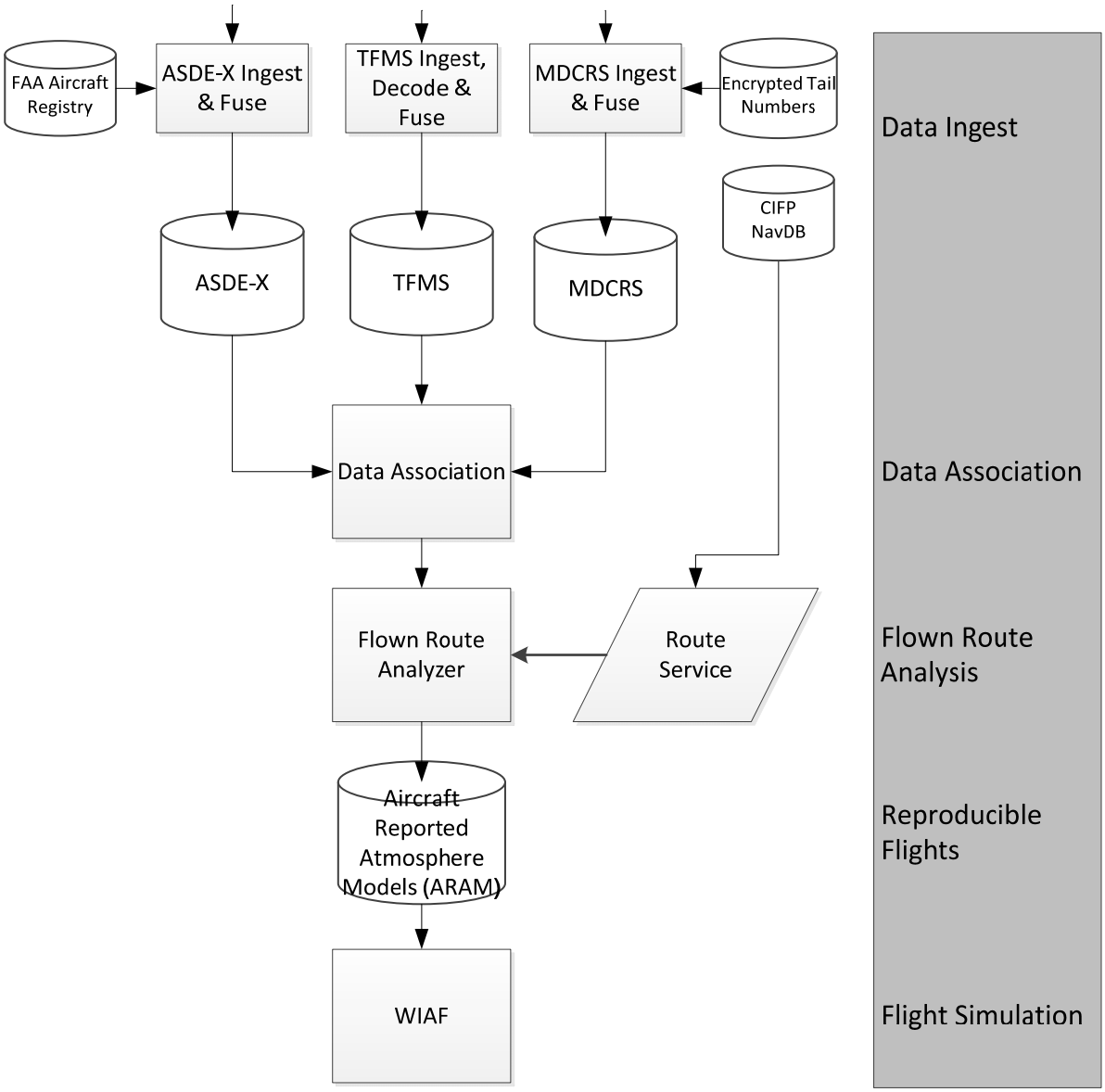

Figure 4. Data processing pipeline to produce correlated (to flight data) aircraft reported atmospheric models (ARAM).

Data Ingest

The first stage of processing is consumption and interpretation of required data sources. The minimum set required is defined below.

- Traffic Flow Management System (TFMS), which provides speed and flight plan information and position updates every minute from origin to destination on domestic flights.

- Airport Surface Detection Equipment-X (ASDE-X), which provides 1 second updates on flights in the terminal area down to the surface, which allows unique aircraft transponder identification.

- Meteorological Data Collection and Reporting System (MDCRS) data, described in Section 3.2.1, was used to provide a basis of truth data for flights. The data source contains information on aircraft position, wind speed and direction, pressure, and temperature. Information is updated every 7 minutes en route and every minute during the descent phase of flight.

- A route data service was constructed based on ARINC 424 descriptions. These descriptions provide a detailed description the route of interest. A number of key fields were used extracted from these descriptions, including the position of each waypoint along the route, speed and altitude constraints on the approach portion of the route, and the location and elevation of the destination airport.

- FAA Aircraft Registry provides the current mapping of aircraft tail number for U.S. registered aircraft to their Mode-S transponder address.

These data are consumed, mostly as streamed data, and processed to fuse individual reports into a comprehensive representation of an aircraft's flight life cycle. These fused data sets are then inserted into separate database tables as elemental components of MAFID.

Data Association

The goal of the second step is to identify the flight information (assigned flight plan, etc.), if any, that is associated with the meteorological data contained in the fused MDCRS data. To do this, we must first identify the tail number of the MDCRS aircraft and then from the tail number determine the particular flight plan it used on that particular flight.

The MDCRS data in the form provided by Meteorological Assimilation Data Ingest System (MADIS) cannot be associated with a particular aircraft because its tail number is encrypted in the MADIS records. With the authorization of Airlines for America (A4A) and the tail number key file provided by the National Oceanic and Atmospheric Administration (NOAA), the fused MDCRS data can be associated with the registered tail number of the physical aircraft.

Next, on the given day of any flight, MAFID-provided ASDE-X records are used that, when processed during ingest, converted the reported Mode-S transponder address to its assigned tail number obtained from the FAA registry database, to look for the tail number of interest. If a match was found, the corresponding flight number used by that commercial aircraft on that part of the day is extracted from the same ASDE-X records.

There is then a link between MDCRS meteorological data during a specific part of a day, the tail number of the airframe, and the flight number it used for that particular flight. The date and time information is then used with the flight number to identify in the TFMS database the assigned flight plan, Standard Terminal Arrival Route (STAR), etc. as specified by ATC.

Flown Route Analyzer

The goal of the second step is to identify those sections of a given flight, if any, where the fused track data indicates that the route flown corresponds to the assigned route as indicated by TFMS flight data. This is necessitated because our intended application of an ARAM is to have simulated aircraft fly identical sections of a route as programmed into the FMS by the TFMS provided route. Though the reported MDCRS data is accurate for the actual trajectory flown, a deviation from the programmed route cannot be modelled without knowing the intent of the pilot or direction given by ATC. In absence of such information, we must sectionalize and preserve only track data that adhere to the assigned route to a high degree. The Route Analyzer tool evaluates this adherence.

The process flow of the tool is depicted in Figure 5. Its architecture is explained in the following summary and description points.

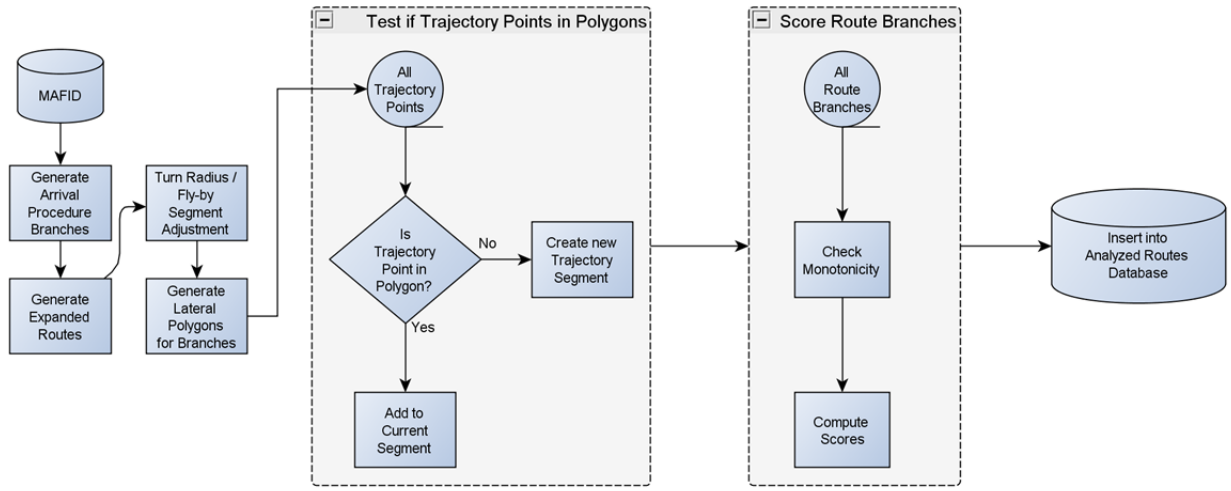

Figure 5. Process flow for analyzing segments of trajectory data that meet routes adherence constraints.

The following steps are used to ensure adherence to flight plan and characterize a particular flight as a candidate for simulation:

1. The assigned route is extracted from the corresponding TFMS data

a. The extracted route is the basis for the analysis, although in itself it does not contain all the data needed to compare a flight against all trajectory permutations during arrival

b. Assigned routes are stored in the TFMS database table, which contains TFMS trajectory and flight plan data spanning multiple months. Assigned routes are actually route summaries that include origin, Standard Instrument Departure (SID), route, STAR, and destination. It does not include the instrument approach procedure (IAP) entered into the FMS, if any, nor the transition into that procedure. Knowledge of these later elements is of particular importance because, for many STARs, this selection defines a significant portion of the waypoints on a given STAR. Here is an example of an assigned route:

    KPHX.FTHLS3.MAXXO..TXO..MLC..MEM.KOLTT1.KATL

c. The above route will be used as part of the analysis to determine if the route was flown without deviation

2. All possible STAR transitions and IAP combinations for the assigned route are computed and stored as route branches that outline descent and approach behavior possibilities

    a. This step expands the analysis field to include valid approach dependent route segments

    b. Since assigned routes do not include the IAP nor IAP transition programmed/flown, these are derived in a combinatorial fashion, constructing a tree in which all feasible combinations are stored as route branches, including combinations that do not use IAP transitions. Some examples of such route branches for approaches into ATL:

        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + JAAJJ ILS08L
        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + LARII ILS08L
        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + ILS08L
        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + GPEAT ILS08R
        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + ILS08R
        KPHX FTHLS3 MAXXO TXO MLC MEM KOLTT1 KATL + AAKAY ILS09L

3. An expanded route is computed for each route branch to generate lists of all flown waypoints

    a. The route and its approach options are expanded to a more detailed level as each route item can be a combination of waypoints. The waypoint lists for each permutation of possible approaches will be used to correlate to the trajectory data collected

    b. A queryable web service has been constructed to perform navigational database lookups. With this service, the expanded route is computed for each route branch to generate all required waypoints.

4. For each route branch, a set of polygons is constructed for evaluation of each trajectory sample, with the width of the polygons accounting for proximity to the destination airport, and additional polygons are constructed around virtual waypoints when turns tighter than 10° are encountered to allow for fly-by behavior

    a. The polygons will ultimately be used to determine if the flown trajectory data points are contained in a valid combination of polygons representing the route options.

    b. From the expanded route branch's lateral coordinates, an altered set of waypoints is computed to account for realistic flight behavior (e.g., fly-by) at each waypoint. If the course angle change between each pair of segments is greater than 10°, a typical turn arc is computed and the waypoint at which the turn occurs is shifted inwards to lie on the arc rather than at the intersection of the two segments. Two additional waypoints, located at the arc insertion coordinates, are added. From this list of altered waypoints, a set of polygons is constructed. The width of the polygons accounts for the radial proximity to the landing runway and is modulated continuously from 4 NM port and starboard (total 8 NM width) down to 1 NM port and starboard (total 2 NM width) See Figure 6. The sigmoid function that defines the polygon width as a function of distance to airport is shown in Figure 7.

5. The flown trajectory points are evaluated against all possible sets of polygons

    a. Each TFMS trajectory (lat, lon) sequence point is evaluated against the polygons that represent all possible route branches.

    b. Continuity of containment in polygons is checked to verify that trajectory points do not step outside the polygon bounds.

    c. Monotonicity of transition between polygons is checked, signifying adherence to a route. A constantly increasing polygon index is a good indicator of correct progression, thus confirming the choice of route branch.

    d. Multiple segments of contiguous polygons may be identified for one flight.

6. Route branch scores are evaluated to select the route branch that best represents the flown trajectory

    a. The TFMS trajectory (lat, lon) sequence is evaluated against the polygons that represent all possible route branches. A scoring N×M matrix is constructed, where N is the number of TFMS trajectory points for a particular flight, and M is the number of possible route branches. For each row N, the (lat, lon) pair representing it is checked for being contained inside one of the polygons for each route branch, and the polygon index is recorded in that matrix element. The scoring matrix is visualized by the heatmap in Figure 8. The "cool" to "hot" progression of color from the top of each column indicates the flight occupying and

sequencing through route segments. Interruptions and missing data appear in solid dark blue and are clearly contrasted against the usable trajectories noted in progressing color.

b. Relative scores are computed for all possible route branches within a scenario using the greatest number of contiguous segments that a route branch steps through, combined with the total number of segments for that route branch, and inversely weighted by the distance of the final usable branch segment to the destination airport. The score produced by this combination of variables has been found to correspond well to the usability of a route branch. In Figure 8, the two red boxes annotate the heatmap, highlighting likely flown branches based on their high scores.

c. Given a selected route branch, its characteristics, including along-track and radial route distances, as well as (lat, lon) point sets, starting and stopping times, and MDCRS data outlined by the segment bounds contained in the route branch, are recorded in an output database table.

7. Contiguous sets of trajectories and their characteristics are stored in database for later consideration

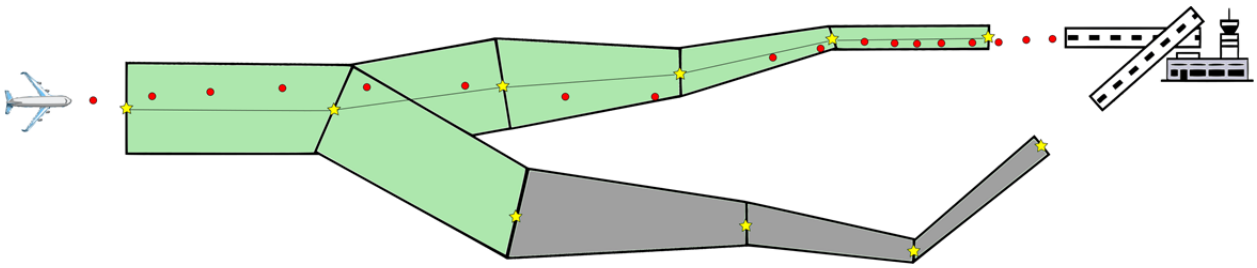

Figure 6. Route containment polygons whose dimensions change as a function of distance from the airport and the incidence angles between waypoints.

The down-selecting from the total of surveyed flights to the qualified candidate flights is presented on a per airport basis in Table 11.

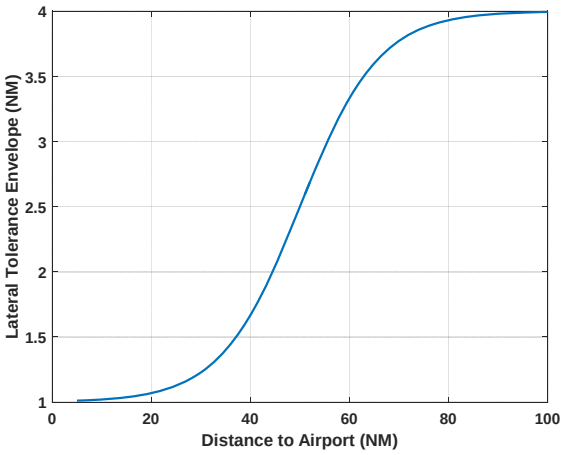

Figure 7. Sigmoid function used to define error tolerance as a function of the track distance from the destination.

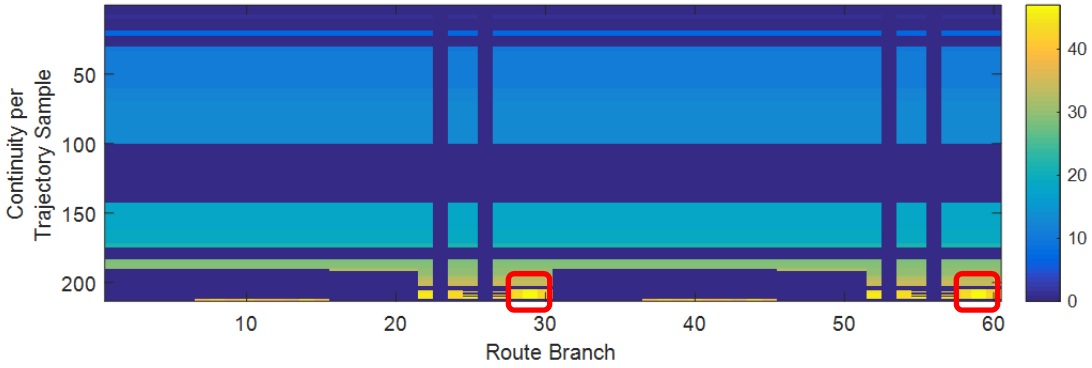

Figure 8. Sample route branching metrics results.

3. ANALYSIS OF WIND INFORMATION

3.1 INTRODUCTION

Previous phases of this work examined the forecast accuracy of operational U.S. numerical weather prediction models used by ATC and airlines' flight planning departments. The models evaluated were the Global Forecast System (GFS), Rapid Refresh (RAP), and High Resolution Rapid Refresh (HRRR). The source of "truth" winds for comparison against the model forecast winds was the HRRR zero-hour analysis winds. The HRRR transitioned from experimental at NOAA Earth System Research Laboratory (NOAA/ESRL) to operational at NOAA/NCEP on September 30, 2014 and represents the state-of-the-art high resolution (3 km) operational numerical model over the CONUS domain. The RMS vector error between the forecasts and the HRRR truth winds was the chosen metric. The RMS vector errors were computed through three-dimensional sampling and aggregational averaging over a 10-month period within four volumes centered over San Francisco International Airport (SFO), Phoenix Sky Harbor International Airport (PHX), Chicago O'Hare International Airport (ORD), and New York Newark Liberty International Airport (EWR).

In Phase 4 of this work, we have performed an extended and complementary assessment of wind forecast model performance wherein the performance has been evaluated along actual flight trajectories using aircraft meteorological data reports from the Meteorological Data Collection and Reporting System (MDCRS) as the source of truth winds for comparisons against the forecasts. We chose to perform these analyses on executed flight trajectories because the results are significantly more operationally relevant from an air traffic perspective than the previous volume-based approach. MDCRS are an appropriate source of truth winds since they are routinely assimilated as observations into the numerical model and have been used frequently by researchers to validate model performance. Since the HRRR represents the latest state-of-the-art U.S. operational model, we focused this assessment on that model, but the GFS is still a commonly used model by ATC and stakeholders, and we plan to extend this trajectory-based analysis to that model in following work.

3.2 DATA SOURCES

3.2.1 MDCRS

MDCRS aircraft observations are collected, processed, and archived by NOAA Global Systems Division (GSD) and are made publically available through the Meteorological Assimilation Data Ingest System (MADIS). More than 100,000 meteorological reports per day from more than 4,000 aircraft are available over the CONUS from the MDCRS system [7]. Figure 9 shows an example of MDCRS data coverage for a single day (February 1, 2016). On this day, there were nearly 125,000 aircraft observations over the CONUS.

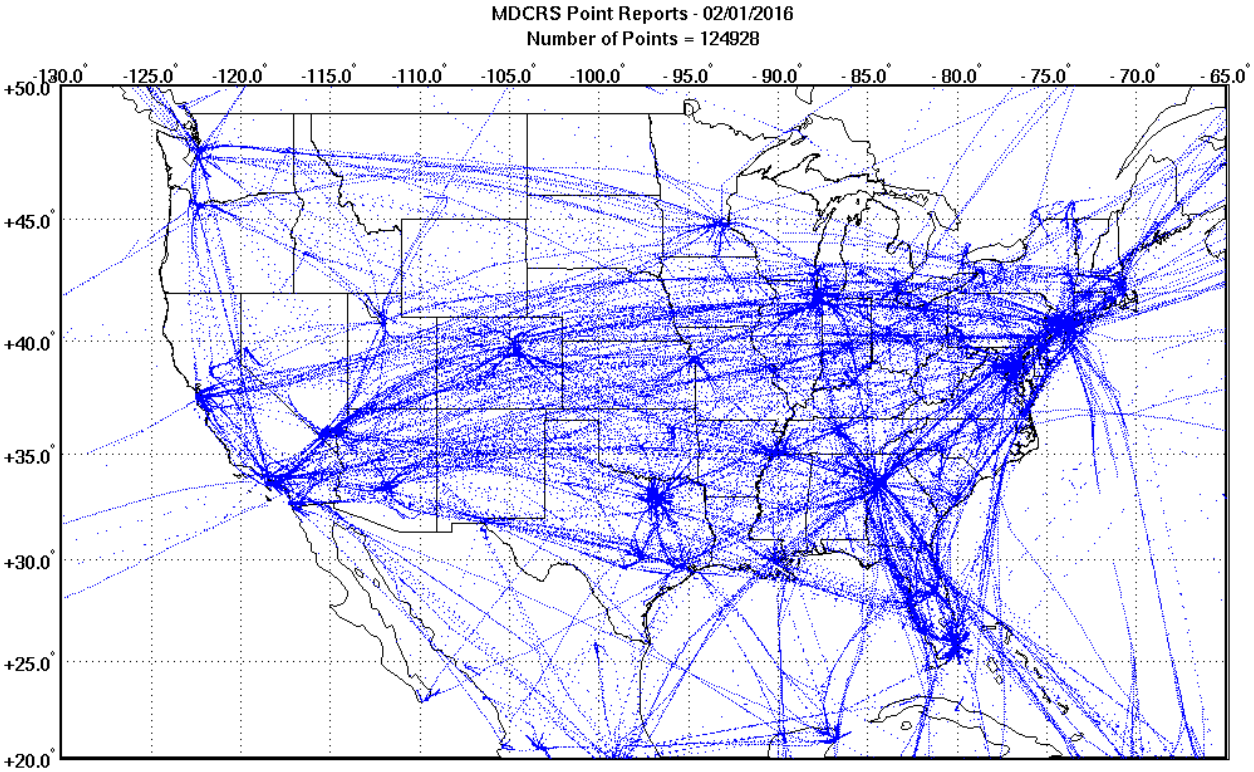

Figure 9. Example MDCRS report coverage over CONUS.

The aircraft reports include position information (latitude, longitude, pressure altitude), time, meteorological variables (temperature, wind speed and direction, etc.), and quality flags for each of the variables. MADIS performs a number of post-processing temporal and positional consistency quality checks prior to storage in their database. Each report contains a "roll flag" field that may or may not be populated, and indicates when the aircraft is maneuvering (turning, banking) such that the associated data report may be unreliable. In our analysis dataset, we found that only 12% of the reports had roll flag information present, with 95% of the reports with available roll flags identified as "good" quality (roll ≤5 degrees), and 5% of the available roll flags as "bad" quality (roll >5 degrees). Given the paucity of available roll flag information, we only rejected aircraft reports on the basis of the roll flag if the information was present and indicated bad roll quality.

Aircraft reports are generally provided at varying nominal update frequencies depending on the phase of flight. Table 2 lists the nominal update frequencies by phase of flight.

TABLE 2

Nominal MDCRS Report Update Rates

Flight Phase	Report Update Rate
Take-Off	6–90 seconds
Departure	20–510 seconds
En-Route	3 minutes if below 465 hPa (~20,000 ft) 7 minutes if above 465 hPa 1 minutes if icing conditions are present
Approach	60 seconds

MADIS aircraft data are provided in two formats: points and profiles. The points data contain all of the available reports without any associations or organization to specific flights. The profile data are a subset of the points data that groups the points by flights for arrivals and departures only, but do not include reports from cruise segments. We found that not all available arrival and departure segments were represented in the MADIS profile data, and reports from cruise altitudes were needed for our study, so we performed our own track associations from the point data using the encrypted tail numbers available with each aircraft report together with spatial and temporal proximity logic. Figure 10 shows 3D and 2D plots of individual points and resulting associated tracks for a single day over the ORD region.

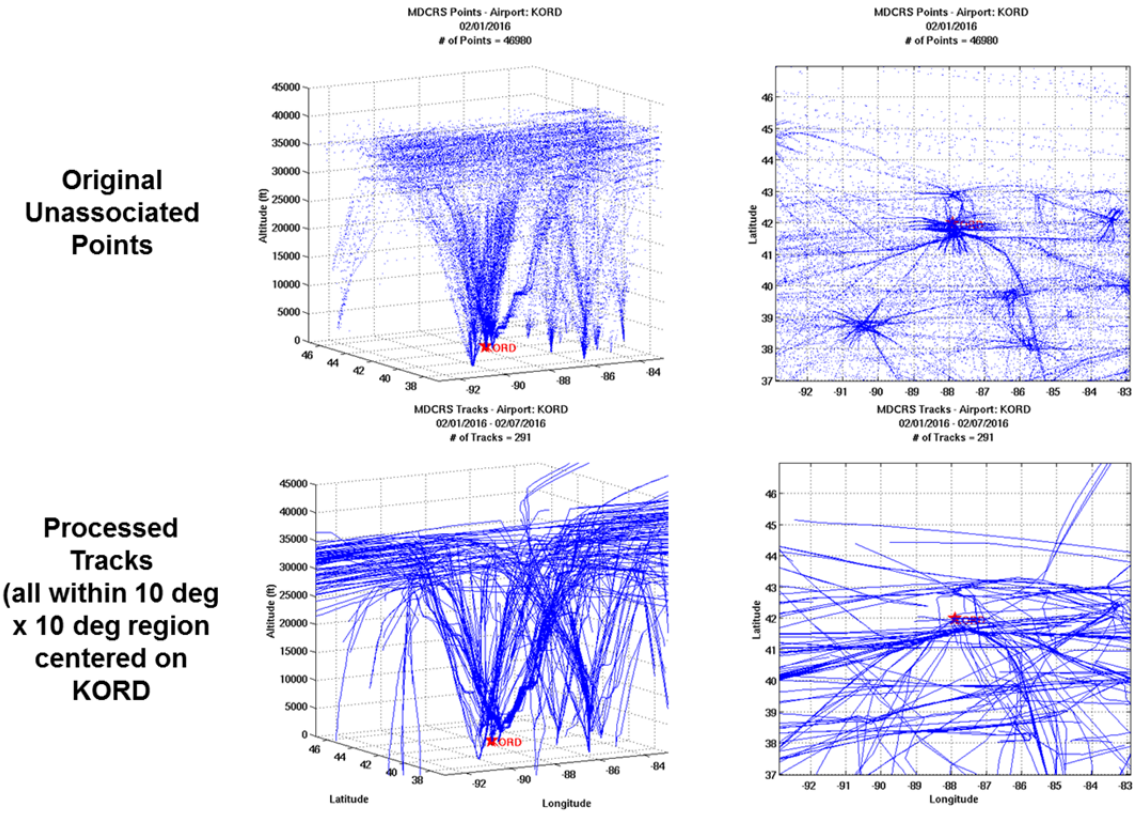

Figure 10. Three-dimensional (left) and two-dimensional (right) views of unassociated MDCRS points (top) and resulting tracks (bottom) after track association. Data are from descents into the ORD region on Feb. 1, 2016.

In addition to storing the position and time information of each MDCRS track sample, track-aggregated wind statistics including wind speed and headwind minima, maxima, means, and standard deviations were computed and stored with the track data in order to support subsequent qualification of tracks based on wind environment characteristics. Headwind values were computed based on along-track changes in direction between successive report locations, not actual aircraft heading information, since that information was typically absent in the MDCRS data. A tail number key file provided by NOAA to Lincoln Laboratory with the authorization of A4A allowed assignment of the actual tail number to each generated track. When combined with other flight information, the actual tail numbers allow look-up and comparisons of the MDCRS-based tracks with filed flight plans.

The processed MDCRS tracks are stored in a database table for additional processing and analysis. As of this writing, the MDCRS tracks database presently contains over 1.6 million tracks from January 1, 2015 through March 31, 2016.

3.2.2 High-Resolution Rapid Refresh (HRRR)

The HRRR model is an hourly updating, 3 km horizontal resolution, CONUS domain numerical weather prediction model produced by NOAA/NCEP. The HRRR updates hourly and provides hourly forecast grid sequences of meteorological variables from 0 to 15 hours. For our analyses, we obtained and processed the pressure vertical coordinate files having 25 hPa vertical resolution extending from 1000 hPa to 50 hPa. Two-dimensional surface variables such as surface pressure, wind, and temperature are also included in the pressure coordinate data files.

3.3 ANALYSIS METHODOLOGY

3.3.1 Geographical Coverage

In order to assess forecast capabilities across different geographic wind environments, HRRR wind forecast comparisons were made over four separate 400 NM × 400 NM regions centered on San Francisco (SFO), Phoenix (PHX), Chicago (ORD), and Newark (EWR) airports. The SFO region provides a relatively "benign" west coast environment dominated by high pressure, low wind speeds, and infrequent wind shear. The PHX region represents an arid climate, and is also a relatively benign wind environment, but has more occurrences of sub-tropical jet stream winds due to its southerly latitude. The ORD region typifies an upper mid-western climate with a good mix of convective summer and winter storms with occasional strong winds and vertical wind shear. The EWR region represents the northeast coastal environment which has a high frequency of strongly sheared wind environments (e.g., Nor'easters), and high winds aloft due to frequent confluence of the polar and subtropical jet streams. The HRRR data were partitioned to these coverage regions as shown in Figure 11.

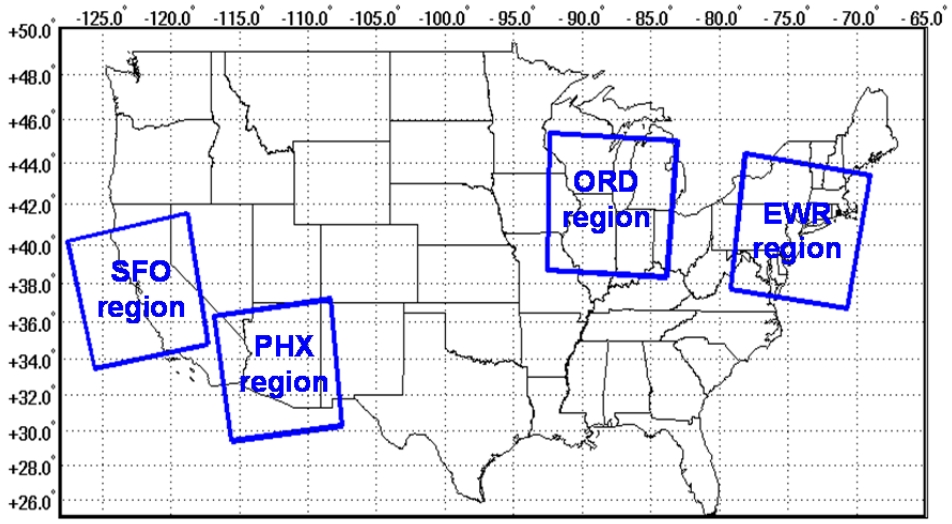

Figure 11. Regions analyzed for Wind Forecast Model performance.

3.3.2 Analysis Period

MDCRS wind observations and matching HRRR wind forecasts from a one-year period extending from March 1, 2015 through February 29, 2016 were compared and analyzed over the four regions for HRRR forecast look-ahead verification times ranging from 0 to 6 hours.

3.3.3 Selection and Matching of MDCRS and HRRR Data

MDCRS Descent Segment Selection

For analysis of wind forecast model performance along descents, MDCRS tracks were first pre-filtered to select candidate tracks having length of at least 200 NM and having a lowest altitude ending point within a 30 NM radius from the airport center and less than 2,000 feet above the airport altitude. Next, the candidate tracks were further reduced to those whose time at the arrival endpoint fell within plus or minus 15 minutes of the HRRR forecast hour. The 15-minute time match criteria was chosen for this and following analyses because it provides an appropriate degree of temporal match given the atmospheric scales of interest without resorting to interpolation between forecast times which would blur the classification of results by the discrete forecast look-head times. Previous researchers [8] have used this approach with time windows larger than this (e.g., 30 minutes), so we feel ours is a conservative match criteria. Finally, an analysis segment for each track was extracted by proceeding backward from the arrival endpoint to the point of intersection of the HRRR data partition region boundary. This resulted in arrival descent track segments of ~200 NM in length. Figure 12 shows an example of selected descent track segments into ORD during the period 11/12/2015 00:08:00 GMT through 11/13/2015 02:32:00 GMT.

MDCRS Cruise Segment Selection

For wind performance analysis along cruise flight segments, MDCRS tracks were first prefiltered to select candidate tracks having length of at least 500 NM, maximum pressure altitude greater than 24,000 feet, and having one or more lateral position coordinates falling within the HRRR analysis region of interest. The initial candidate tracks were then time-filtered to select those tracks whose start and end times spanned the selected comparison HRRR forecast hour. The final sequence of cruise analysis segment points were extracted by selecting only those points whose report times fell within 15 minutes of the HRRR forecast valid time hour, pressure altitudes were greater than 24,000 feet.

Matching of HRRR and MDCRS Data

HRRR forecasts were matched horizontally to the latitude and longitude of the MDCRS observations using bilinear interpolation. Vertical interpolation of HRRR data to the MDCRS pressure altitudes was performed by first converting the MDCRS pressure altitudes to their equivalent U.S. Standard Atmosphere pressures and using linear interpolation in log(p) across the 25 hPa increment HRRR pressure surfaces. Temporally, MDCRS observations were matched to HRRR forecasts having valid times within plus or minus 15 minutes of the MDCRS observation – there was no interpolation between HRRR forecasts. This facilitates categorization of performance results by the discrete hourly

forecast look-ahead times, while providing a reasonable temporal match given the atmospheric temporal scale of concern.

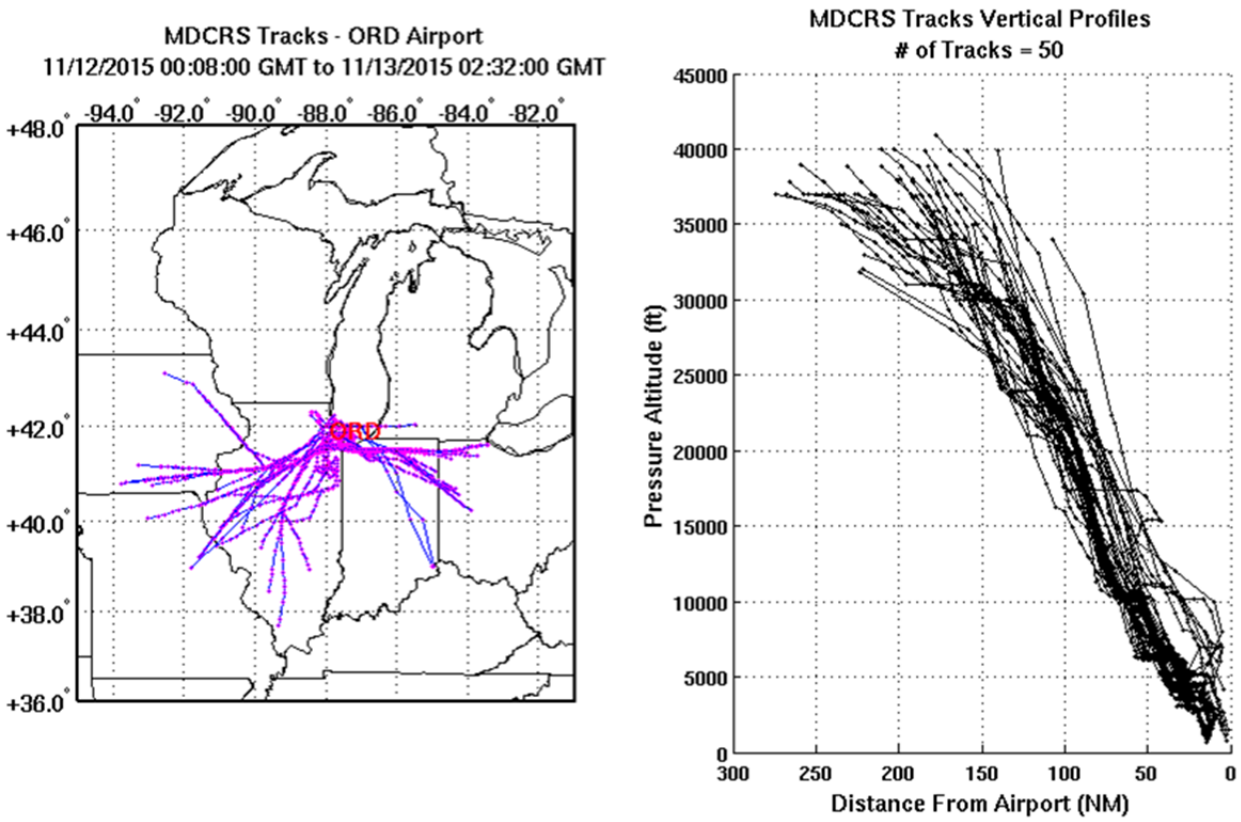

Figure 12. Example of MDCRS descent track segments. Left plot shows lateral track segment locations. Right plot shows vertical profiles of track segments.

3.3.4 Categorization of Results

Wind forecast model performance discussed in the following sections is categorized by:

- Forecast look-ahead time (0,1,2,3,4,5, and 6 hours)
- Flight phase (descent, cruise)
- Wind magnitude
- Altitude
- Location

- Time of year
- Time of day

3.4 METRICS

The following metrics were computed along actual aircraft trajectories and used to assess the effects of wind information quality on 4D-TBO performance in Phase 4 of this work.

3.4.1 RMS Vector Difference (RMSVD)

The Root Mean Square Vector Difference is the root mean square difference applied to the magnitudes of the forecast and observed wind vector components as given by Equation (1):

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

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. It is one of the most commonly used metrics to quantify performance of wind forecast models.

3.4.2 Mean Absolute Error (MAE)

The Mean Absolute Error is the average of the absolute value of the difference between the forecast (f) and the observation (o) as given by Equation (2). This metric weights positive and negative errors equally, making it a measure of total forecast error and is used in this study for quantifying headwind component differences between HRRR forecasts and MDCRS observations.

$$MAE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|f_n - o_n|} \tag{2}$$

3.4.3 Estimated Time-to-Fly (ETTF) Difference

This metric compares the integrated effects of actual (MDCRS) vs. forecast (HRRR) headwind differences in the resulting time-to-fly along analyzed MDCRS trajectory segments and is directly relatable to 4D-TBO time-based targets. The headwinds for each wind source were first integrated using 5-second step interpolation along MDCRS trajectory segments that intersected the available regions of HRRR forecast wind data. The estimated time-to-fly (ETTF) for a trajectory was then computed using

$$\text{ETTF} = \sum_{i=0}^{n} \Delta t_i \tag{3}$$

where $\Delta t_i = \Delta d_i / (\text{airspeed}_i - \text{headwind}_i)$, and Δd_i = length of a step along the flight trajectory.

Airspeeds were estimated by adding the actual or forecast headwinds at the MDCRS trajectory points to the aircraft ground speeds (determined from time and distance traveled between successive MDCRS trajectory points).

3.4.4 Outlier Rejection

Although NOAA MADIS imposes a number of quality control factors in their postprocessing of the MDCRS reports, anomalous reports can still occur due to malfunctioning aircraft sensors, problems in data recording or transmission, etc. When repeated anomalous wind observations are noticed for a given MDCRS tail number, we flagged the aircraft as "bad," and excluded any reports from that aircraft for the analyses. We found that the RMS vector difference between the aircraft reports and the HRRR forecasts to be a useful metric for rejecting outliers. We chose a conservative value of 200 knots as the threshold for outlier rejection, since there could be actual very large forecast differences in high gradient wind environments such as jet stream boundaries.

3.5 RESULTS

3.5.1 Wind Distributions

Figure 13 presents scatter plots of the wind speeds and directions obtained from the MDCRS reports for the four regions that were analyzed for this study over the one-year period of March 1, 2015 through February 29, 2016. The plots show that a good distribution of wind speeds and directions are present in the analyzed data set, and that the higher wind speeds are typically associated with westerly to northwesterly winds.

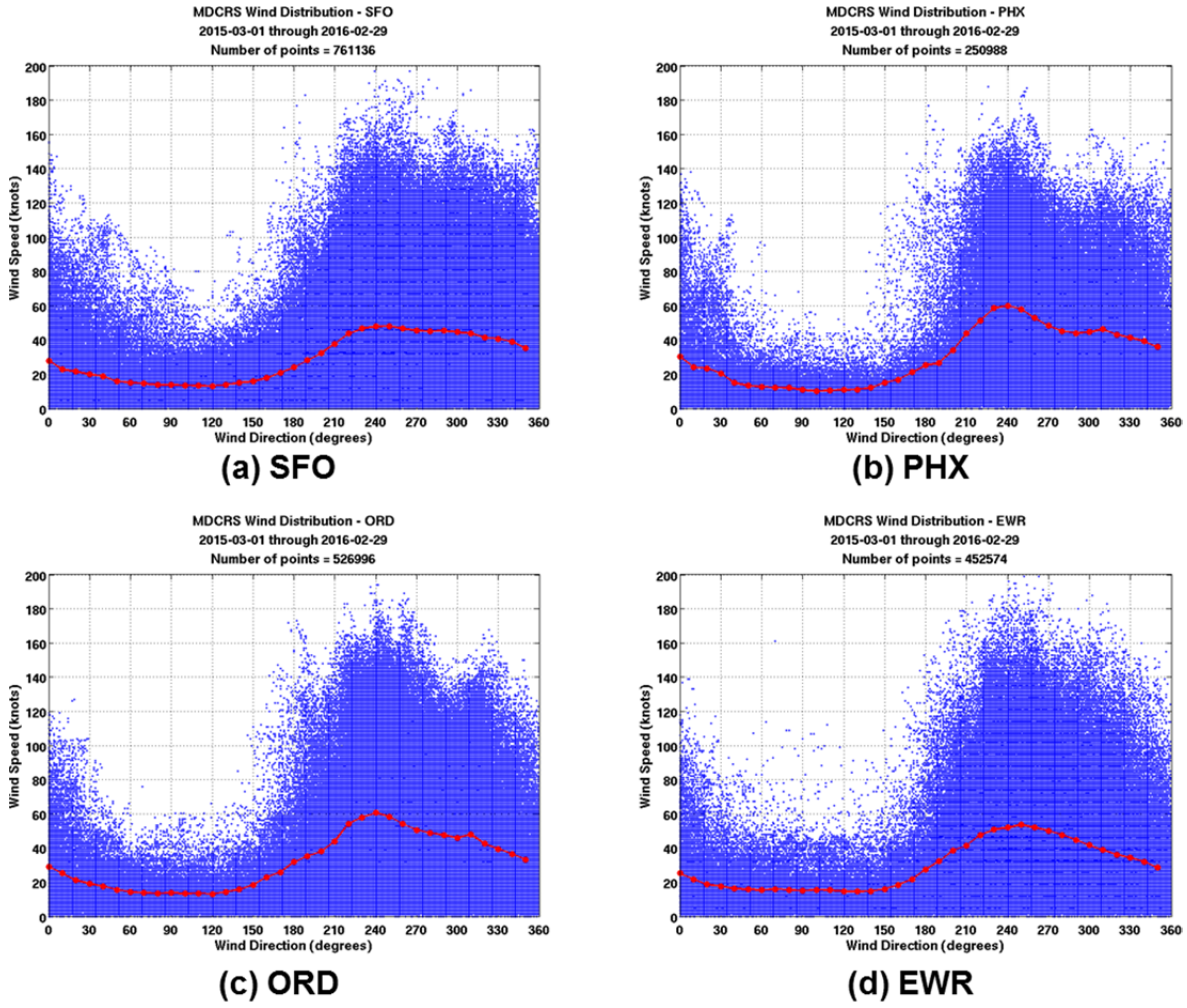

Figure 13. MDCRS observed wind speeds and directions for (a) SFO, (b) PHX, (c) ORD, and (d) EWR. Red points and lines indicate mean wind speed.

3.5.2 Wind Forecast Accuracy by Altitude and Look-Ahead Time

Figure 14 plots the means and means plus one standard deviation of the headwind differences between HRRR forecast and MDCRS observations as a function of altitude and forecast look-ahead time across the four regions over the one-year analysis period. The statistics were binned at 100 hPa altitude intervals with each point plotted at the lower point of the averaging bin. For example, a point plotted at the 1000 hPa level reflects the statistic computed from the differences aggregated over the 900 hPa to 1000 hPa vertical layer.

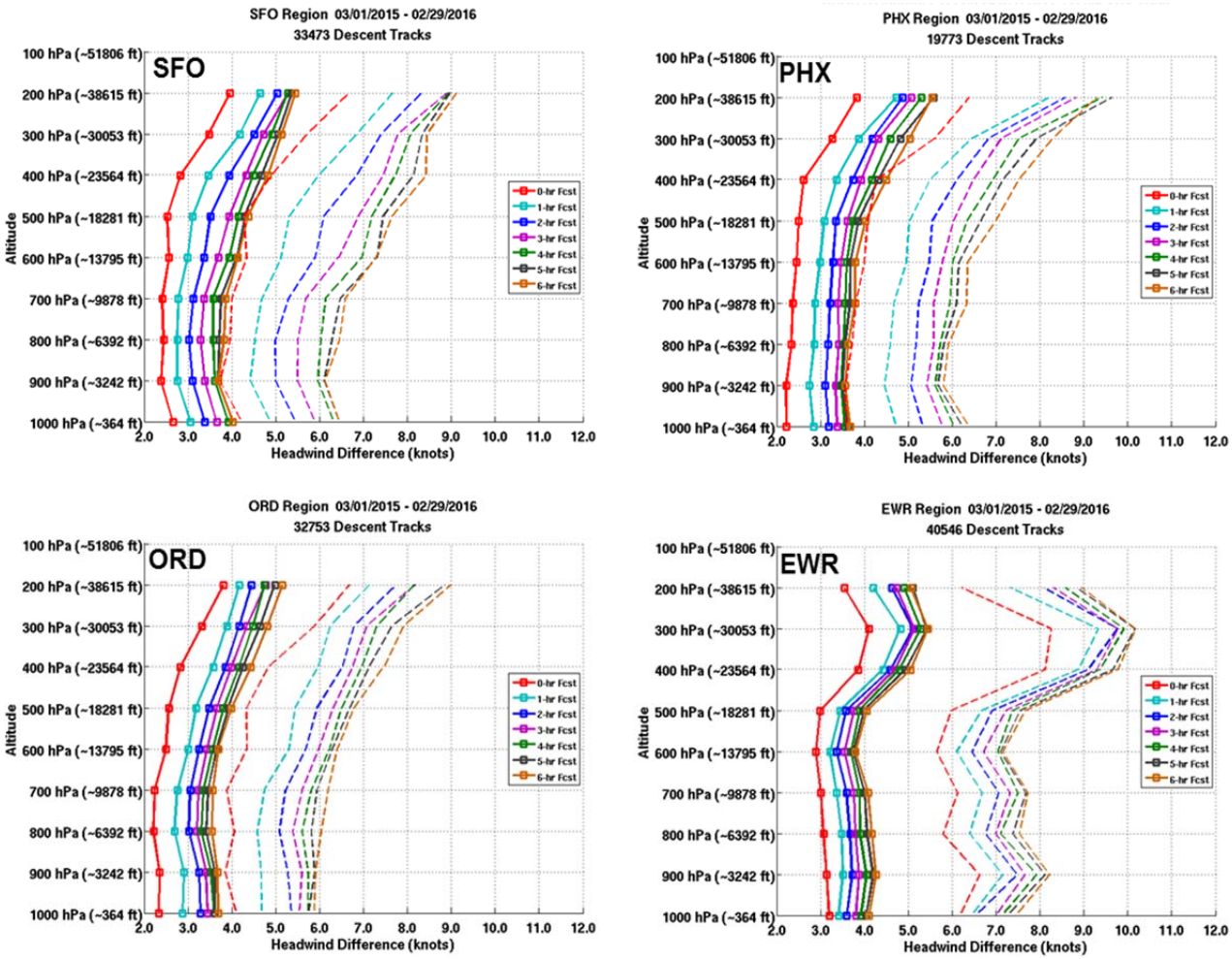

Figure 14. Mean headwind difference (solid curves) and mean plus one standard deviation (dashed curves) between HRRR forecasts and MDCRS wind observations by altitude and forecast look-ahead time for SFO, PHX, ORD, and EWR airport regions.

The forecast differences are seen to generally increase with forecast look-ahead time and with altitude. Mean headwind differences between HRRR forecasts and MDCRS observations ranged from 2.2 knots for the HRRR zero-hour analysis to 5.6 knots for the 6-hour HRRR forecast. Standard deviations ranged from 1.3 knots to 4.7 knots.

The curves are quite similar for SFO, PHX, and ORD regions, but EWR has notably different altitude variation of the forecast differences than the other three regions. EWR exhibits a nearly constant amount of mean forecast differences along each look-ahead time from the surface to approximately 18,000 feet, and then has marked increases in forecast differences from 18,000 to 30,000 feet, followed

by decreases above 30,000 feet. One possible explanation for this is climatological. The Northeast has one of the highest frequencies of jet stream crossings in the U.S. due to frequent confluence of the polar and subtropical jet streams in this region of the country. The jet stream typically occurs at altitudes near 30,000 feet and typically has strong wind gradients that could result in larger wind forecast errors if the model fails to forecast the jet stream location and timing precisely. As has been shown in prior studies [8],[9] and in the following section of this study, numerical model wind forecast errors tend to increase with increasing wind speeds.

3.5.3 Wind Forecast Accuracy by Wind Magnitude

This section examines the relationship of headwind forecast accuracy to wind magnitude. Means (solid curves) and standard deviations (dashed curves) of the differences between the HRRR forecast headwinds and the MDCRs observed headwinds as functions of the MDCRS observed wind speeds are presented in Figure 15. The results are shown for forecast look-aheads of 0–6 hours and for wind speed intervals of <20 knots, 20–40 knots, 40–60 knots, and ≥80 knots. Generally, as wind speeds increased, the forecast differences and standard deviations also increased. The difference trends appear to level off or slightly decrease for wind speeds greater than or equal to 80 knots, but this may be due to fewer samples in this wind speed category (a few hundred samples versus several thousand for the other wind speed categories).

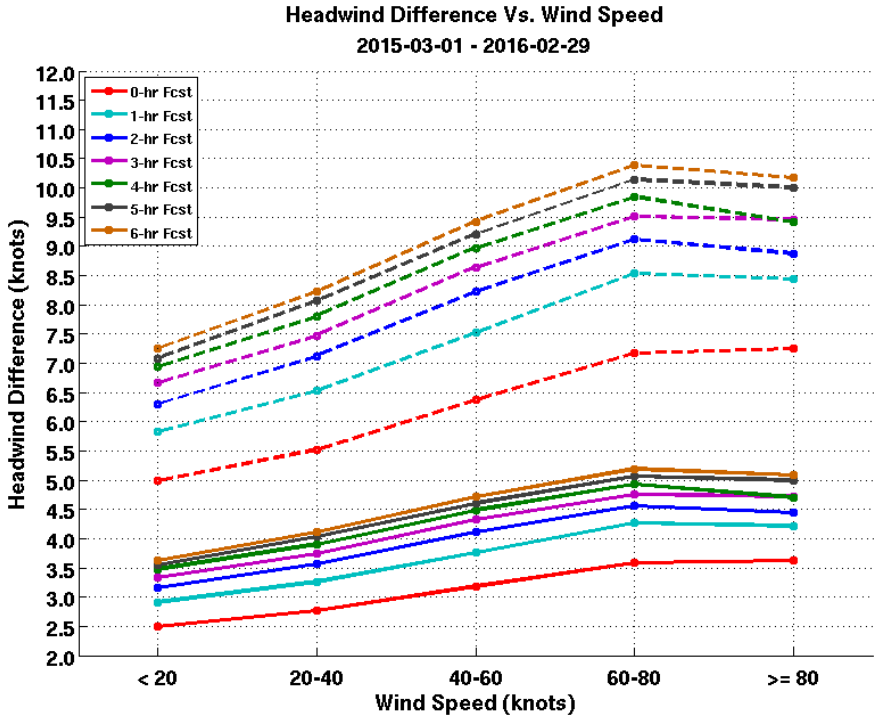

Figure 15. Mean headwind forecast difference (solid curves) and mean plus one standard deviation (dashed curves) by forecast look-ahead time and wind speed.

3.5.4 Wind Forecast Accuracy by Phase of Flight

This section examines the relationship of forecast accuracy to two phases of flight: cruise and descent. For this evaluation, cruise segments from MDCRS reports during the month of February 2016 crossing over SFO, PHX, ORD, and EWR regions were examined. To qualify as part of a cruise segment, the MDCRS pressure altitude had to be greater than 24,000 feet. Figure 16 and Figure 17 plot the mean absolute headwind differences between HRRR forecasts and MDCRS observations for cruise and descent phases, respectively (mean headwind differences were averaged across all altitudes for the descent data). Mean headwind differences and standard deviations were found to be generally 1 to 1.5 knots larger with the cruise segments compared to the descent segments. This is likely due to the higher wind speeds and higher wind gradients at cruise altitudes where the jet stream winds are more likely to be encountered. As seen in the previous examination of wind speed effects, wind forecast accuracy tends to decrease in higher wind speed environments.

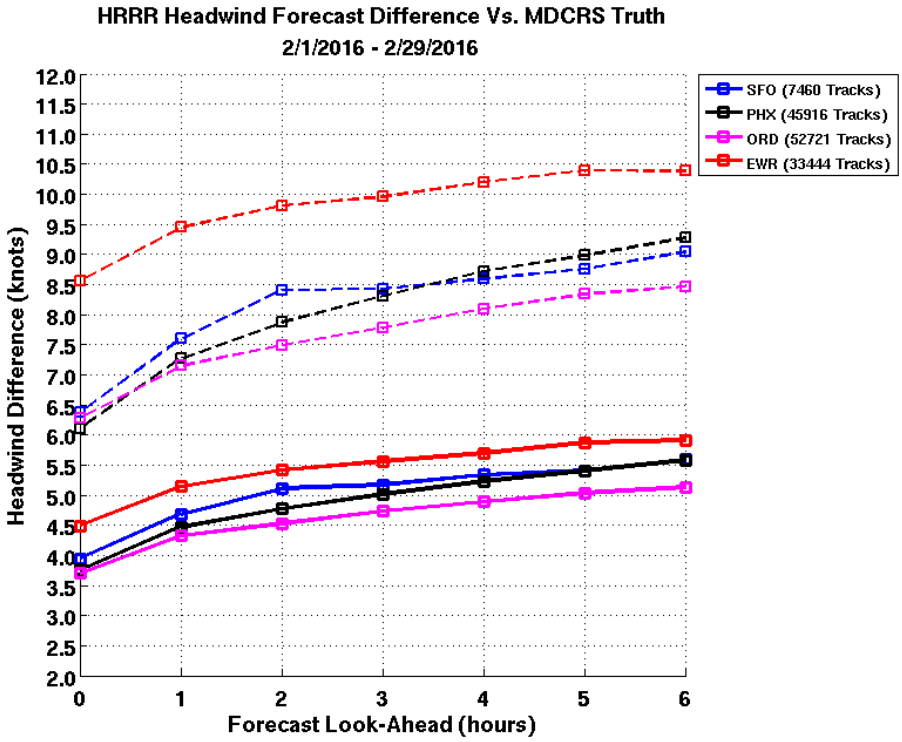

Figure 16. Mean absolute headwind difference (solid curves) and mean plus one standard deviation (dashed curves) for cruise flight segments.

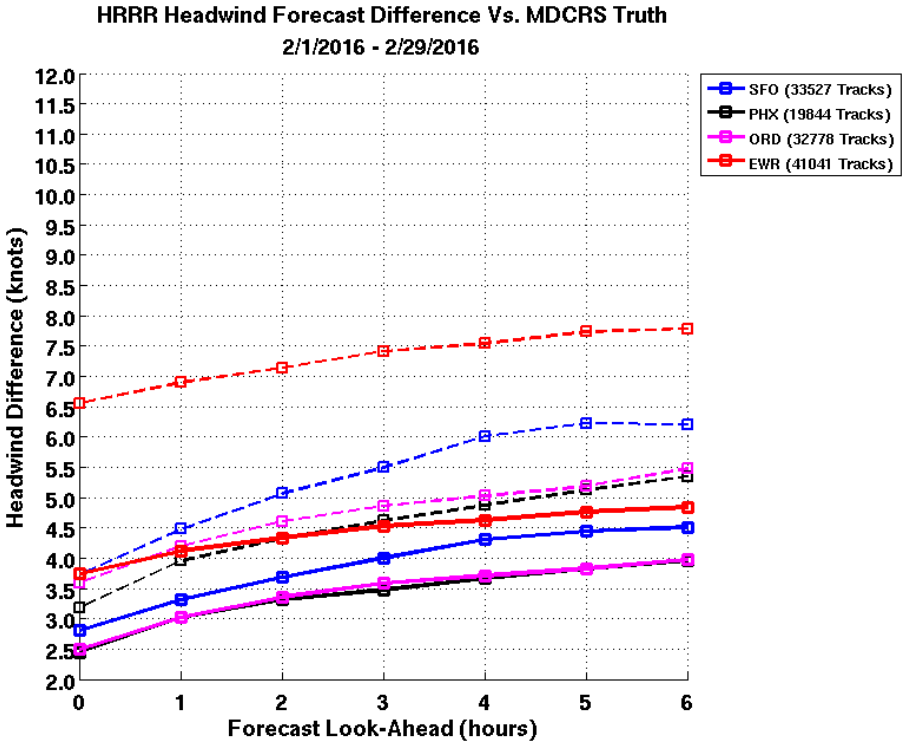

Figure 17. Mean absolute headwind difference (solid curves) and mean plus one standard deviation (dashed curves) for descent flight segments.

3.5.5 Wind Forecast Accuracy by Time of Year

In this section, we examine the effects of time of year on wind forecast accuracy. Headwind forecasts based on the three-hour and six-hour HRRR forecasts were compared against the MDCRS headwinds. The results, aggregated across all four of the analyses regions and over all altitudes, are shown in Figure 18 and Figure 19. As expected, mean headwind forecast differences and standard deviations were found to be larger in the winter months for both forecast lead times, consistent with higher wind speeds and gradients in the winter. The seasonal effect is most pronounced for the EWR region, presumably owing to the climatologically larger occurrence of high winds from coastal storms and jet streams overhead than the other sites in the winter months.

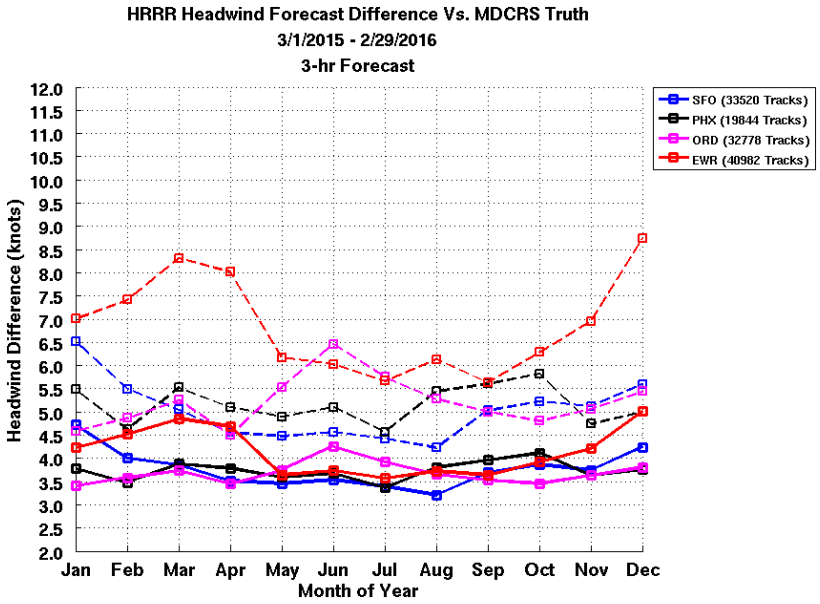

Figure 18. Monthly mean (solid curves) and mean plus one standard deviation (dashed curves) headwind forecast differences between 3-hour HRRR forecasts and MDCRS observations.

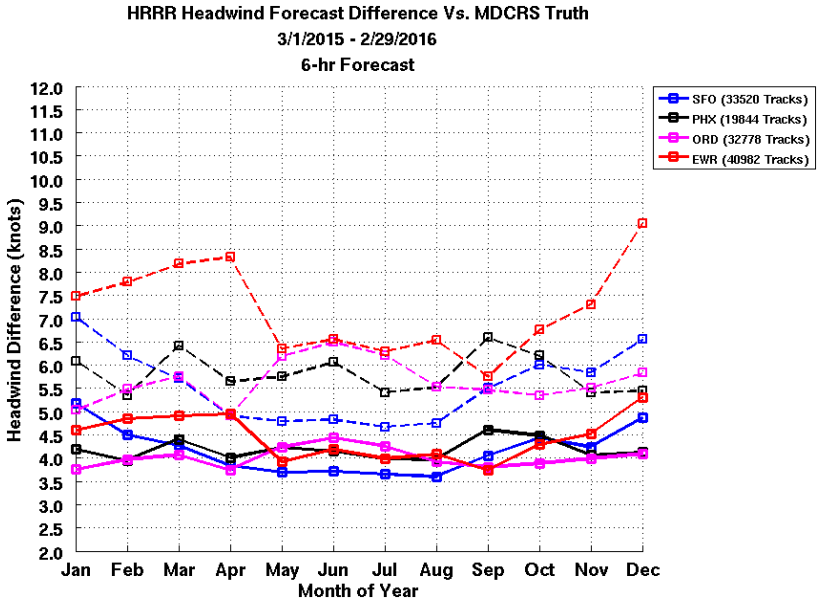

Figure 19. Monthly mean (solid curves) and mean plus one standard deviation (dashed curves) headwind forecast differences between 6-hour HRRR forecasts and MDCRS observations.

3.5.6 Wind Forecast Accuracy by Time of Day

This section examines the effects of time of day on wind forecast accuracy. Headwind forecasts based on the 3-hour and 6-hour HRRR forecasts were compared against the MDCRS headwinds at each of 24 hours. The results, aggregated across all four of the analyses regions and over all altitudes, are shown in Figure 20 and Figure 21. Mean headwind differences for the two forecast look-aheads ranged between 3.3 knots and 5.4 knots. The EWR region exhibited slightly higher mean differences between 9:00 a.m. and 12:00 p.m. Standard deviations of the forecast differences were markedly higher for EWR during the same hours of the day, and remain higher than the other sites for the later hours of the day.

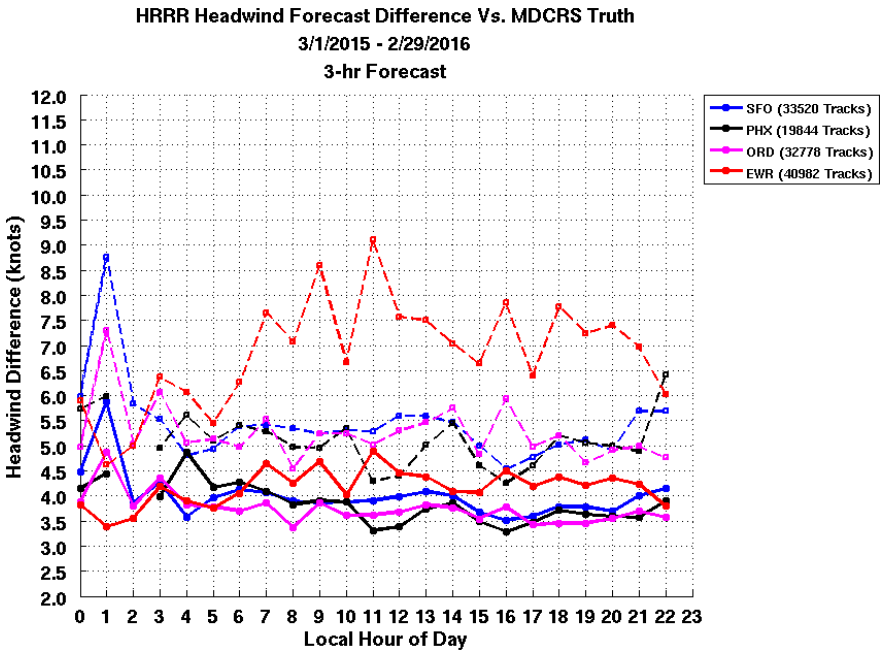

Figure 20. Mean (solid curves) and mean plus one standard deviation (dashed curves) of headwind forecast differences versus time of day (Local Standard Time) between 3-hour HRRR forecasts and MDCRS observations.

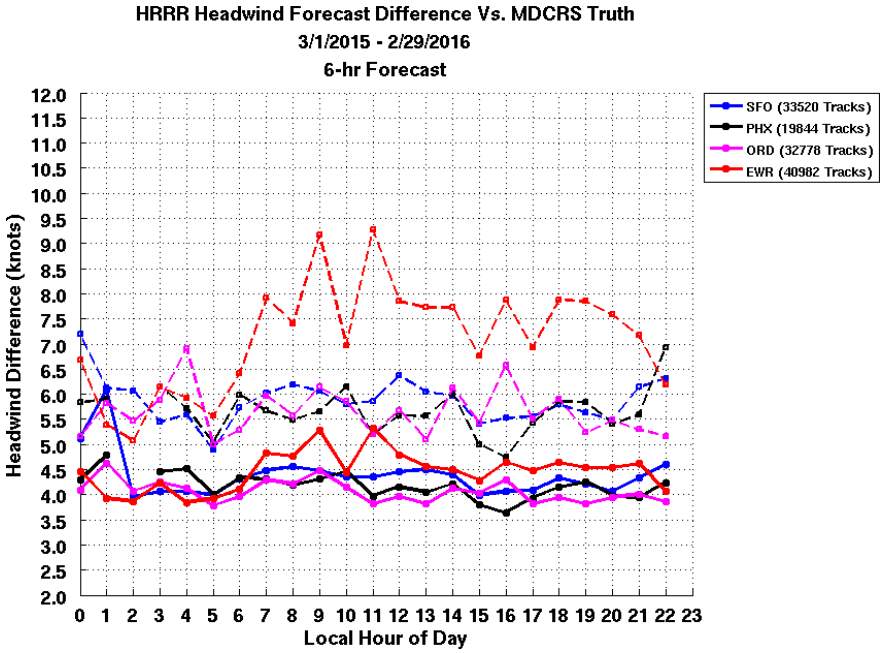

Figure 21. Mean (solid curves) and mean plus one standard deviation (dashed curves) of headwind forecast differences versus time of day (Local Standard Time) between 6-hour HRRR forecasts and MDCRS observations.

3.5.7 Comparison of Headwind Differences and RMS Vector Differences

While this phase of the research has focused on trajectory-based analyses and the use of headwind differences between HRRR forecasts and MDCRS observations as the principal metric, it is informative to relate this metric to the RMS Vector Difference (RMSVD) metric commonly reported in wind forecast model research literature and in prior phases of this research. From purely geometrical considerations, if the analyzed samples have wind directions equally distributed throughout the compass, RMS vector differences would be expected to be approximately 1.4 (i.e., sqrt(2)) times larger than headwind component differences computed for the same samples. Prior research [8] has noted that the RMS vector difference metric is biased in high wind speeds, where a small wind direction difference between a forecast and observation can result in a large vector difference.

Figure 22 plots the mean and standard deviation headwind component differences between HRRR forecasts and MDCRS observations averaged across all altitudes for each of the four regions as a function of forecast look-ahead. Figure 23 compares the same forecasts and observations using the RMS vector differences. The headwind differences range from approximately 2.5 to 4.5 knots, while the RMS vector differences range from approximately 4.5 to 8.0 knots, which is more than the amount that would be ascribed to purely geometrical considerations between the two metrics alone. Winds over the U.S. tend to be westerly (strongest winds tend to be from the northwest as shown in the wind distributions previously), so this wind direction bias, coupled with the directions of the specific MDCRS flight tracks and the sensitivity of the RMSVD metric to high wind speeds may account for the additional differences.

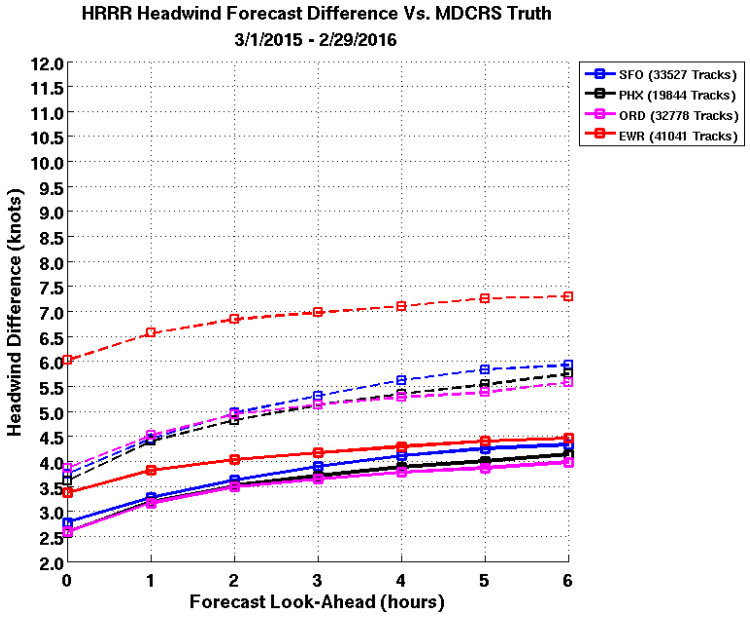

Figure 22. Mean absolute headwind difference (solid curves) and mean plus one standard deviation (dashed curves) between HRRR forecasts and MDCRS observations averaged across all altitudes.

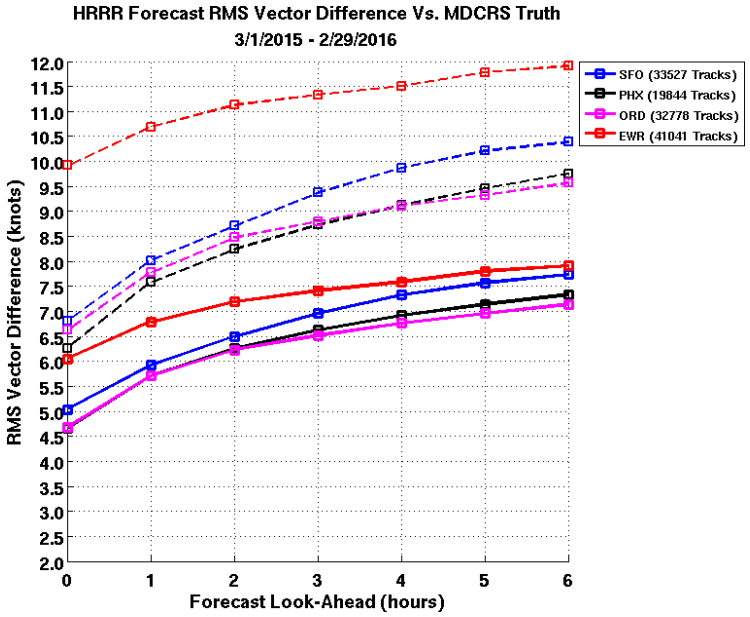

Figure 23. Mean RMS vector difference (solid curves) and mean plus one standard deviation (dashed curves) between HRRR forecasts and MDCRS observations averaged across all altitudes.

3.5.8 Estimated Time-to-Fly (ETTF) Difference

In this section, we examine the effects of differing forecast winds on ETTF. This is an important metric for understanding the effects of wind forecast errors on 4D-TBO procedures such as RTA and IM, as well as other time-based decision support tools that must estimate the time it takes an aircraft to fly along a given path. A similar analysis is described in [10], which used assumed nominal flight plans into Dallas/Fort Worth International Airport (KDFW), and modeled step-wise descent airspeeds. Our analysis uses actual flight trajectories from the MDCRS reports and actual measured time-to-fly times for comparisons against the ETTF, which is what the authors recommended.

We compared the ETTF computed from HRRR forecast headwinds as described in Section 3.4.3 against the actual time-to-fly (ATTF) obtained from the MDCRS tracks. To ensure an adequate number of samples, we imposed a minimum MDCRS trajectory segment duration of 15 minutes for computing the ETTF differences. Because the qualifying segments are of varying lengths, the ETTF differences are normalized by one minute of flight time to allow inter-comparison between results of different flight segments and across different forecast look-ahead times, regions, etc.

Figure 24 plots the means and one standard deviation error bars of the 1-minute normalized ETTF differences between the different HRRR forecasts' look-ahead times and the ATTF obtained from the MDCRS tracks. The mean ETTF differences across all sites were found to be very similar. The ORD region had the largest mean ETTF differences approaching 0.2 seconds per minute of flight. This would translate to an average 6 seconds of time-to-fly difference for a 30-minute flight. Variability of the ETTF difference appears quite high with standard deviations running generally from 0.4 to 0.6 seconds per minute of flight, which translates to 12–18 seconds for 30 minutes of flight.

EWR ETTF difference standard deviations were notably higher, with a standard deviation of nearly 7 seconds at the 6-hour forecast look-ahead (note the different y-axis scale for the EWR results). This would translate to a 3.5-minute ETTF difference for a 30-minute flight. One possible explanation for the large variance is that there are MDCRS report outliers that were not excluded by the conservative 200 knot RMS vector difference outlier threshold test (one of the MDCRS tracks examined had RMS vector differences from HRRR forecasts of 170 knots). In addition, we found some large ETTF difference outliers associated with MDCRS tracks having a relatively sparse number of MDCRS reports in combination with relatively large HRRR headwind forecast errors at those report locations. These forecast headwind samples were then propagated over the large sample distances during the ETTF computational integration, leading to large ETTF differences compared to the MDCRS-based ETTFs. For the following phase, we plan to rework the ETTF difference analysis with additional quality constraints imposed on the associated numerical model forecasts and on the trajectories selected for the analysis to ensure there is sufficient MDCRS sample point density to support the computations.

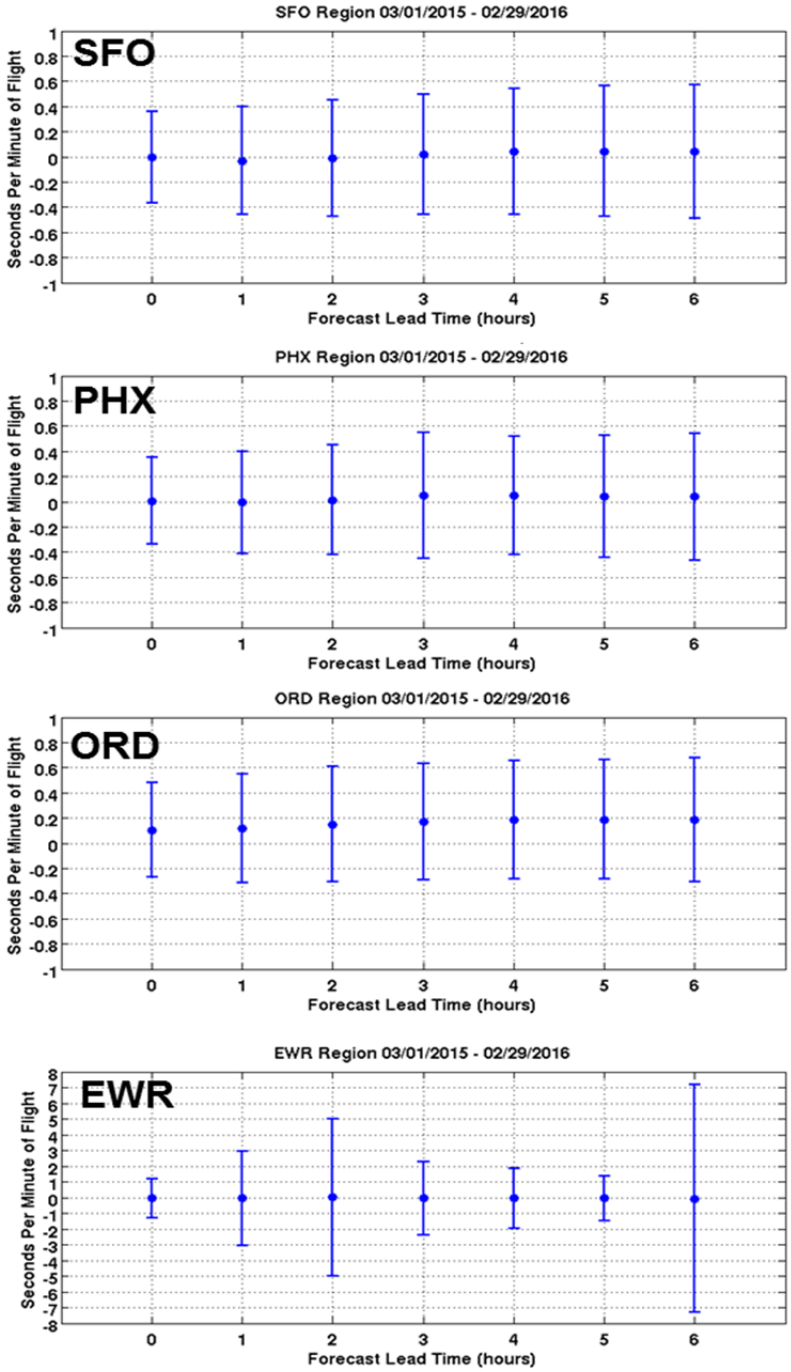

Figure 24. Means and standard deviations of normalized ETTF difference per minute of flight between HRRR forecasts and MDCRS for the four analyzed regions.

3.6 KEY TAKE-AWAYS

The following summarizes the key accomplishments and findings of the wind information analysis conducted in Phase 4 of this research:

1. Trajectory-based analyses of wind forecast model performance was conducted using aircraft wind observations from MDCRS as the truth wind source for comparison against HRRR model forecasts over a one-year period and four airport regions. MDCRS was chosen as the source of wind truth for this study because it provides directly sensed observations of the winds, is less smoothed than model-based analysis winds, and more precisely captures "flight weather" along actual flown tracks.

2. In addition to the RMS Vector Error (RMSVE) metric used in our previous studies (recast in this study phase as RMS Vector Difference), the trajectory-based analyses in this study permitted evaluation using additional performance metrics such as Mean Absolute Error (MAE) applied to scalar forecast headwind component differences, and ETTF differences that arise from the accumulated ground speed effects of time-integrated headwind differences between different forecasts. The ETTF metric is particularly relatable to time-based targets such as RTA and IM.

3. Results were stratified by a number of different potentially influencing criteria including altitude, forecast look-ahead time, wind magnitude, phase of flight, time of year, and time of day.

4. Mean HRRR headwind forecast differences compared to MDCRS were found to generally increase with increasing forecast look-ahead time and with altitude for all four regions ranging from 2.2 knots to 5.6 knots. Standard deviations ranged from 1.3 to 4.7 knots. Forecast differences were also shown to increase with increasing wind speed, and in the winter months. Only the EWR region showed appreciable changes in mean forecast accuracy with time of day. Mean forecast differences were found to be 1.0–1.5 knots greater for cruise versus descent phases of flight.

5. Mean forecast ETTF differences compared to actual MDCRS flight times across the four airport regions were found to be similar. ORD had the largest mean ETTF differences, with an equivalent 6-minute ETTF difference for a nominal 30-minute flight at forecast look-aheads of 2–6 hours. Variability of the forecast-based ETTFs was quite high, with standard deviations generally ranging from 12–18 minutes for a nominal 30-minute flight. However, EWR had notably much higher standard deviations, especially at the 6-hour look-ahead time.

6. Regionally, EWR exhibited the largest variability in wind forecast accuracy across the various result categories, sometimes dramatically so. It is hypothesized that this is at least in part due to

the climatologically dynamic wind environment of this region, but more investigation is needed to determine if there are other contributing causes, including occasional bad aircraft reports.

3.7 RECOMMENDED NEXT STEPS

The following are some recommended next steps for further analysis and understanding of wind forecast model performance:

1. Extend the MDCRS truth and trajectory-based wind performance analysis to the GFS model, which is a commonly used operational model by ATC and stakeholders.

2. Quantify frequencies of occurrence, durations, and wind environment characteristics (e.g., vertical wind shear) associated with large wind forecast errors. Special attention should be given to the EWR region, since relatively large error standard deviations were noted for this region in the Phase 4 results. The New York region is also a high-traffic airport region where operational impacts of poor wind forecasts on 4D-TBO procedures would be potentially significant.

3. Explore the ability to predict ahead of time expected wind forecast model performance, e.g., over the next 24 hours, how good are the various forecast models expected to perform? Research is needed to determine if problematic forecast wind environments can be automatically identified (e.g., from combination of real-time wind environment analysis and monitoring of recent model forecast performance) and used to provide decision support for the feasibility of successfully executing a given wind information-dependent procedure.

4. Perform comparisons of wind model forecasts against high-update Mode-S EHS aircraft reports. Comparisons against these high-resolution wind observations may yield insight into wind changes not captured by the forecast models.

5. Explore uses and benefits of a real-time MDCRS (or comparable system) aircraft meteorological data feed. Lincoln Laboratory could conduct flight trials to support this research if appropriate.

4. RTCA SPECIAL COMMITTEE ACTIVITIES

4.1 RTCA SC-206 SUPPORT

As part of Phase 3 work, Lincoln Laboratory (LL) was initially tasked to provide a limited amount of support to RTCA Special Committee 206 (SC-206), "Aeronautical Information Services (AIS) and Meteorological Data Link Services," and other stakeholders interested in the issue of wind information quality in air transportation operations. This tasking was greatly expanded under the direction of the sponsor and included the appointment of LL as a co-chair of a newly formed subgroup of SC-206 denoted as SG-7. The role of SG-7 based on the RTCA Terms of Reference is to "Develop Guidance for the Use of Data Linked Current and Forecast Wind Information in Air Traffic Management (ATM)." This subgroup must consider the effects of wind information in terms of quality and availability necessary to support a number of NextGen programs including Interval Management, 4D-TBO (e.g., required time of arrival (RTA)) and wake vortex mitigation operations.

In addition to the organizational and administrative roles as chair of the sub group, LL was directed by sponsorship to concentrate on IM and RTA related research and analyses, as these programs are directly related to Weather Technology in the Cockpit (WTIC) research areas.

One of the principal deliverables from LL to SG-7 was a set of hypotheses to be considered during the September 2015 SC-206 plenary session regarding performance questions related to IM and RTA operations. These initial hypotheses were considered and honed by SG-7 plenary attendees and members whose expertise covered the domains of ATC, airline piloting, aviation meteorology, airline operations, avionics, and general IM and RTA research.

The original high-level hypotheses have been slightly refined since then and have been used to develop a test matrix and test plan that, once executed, would provide material data that could be used in support of the development of the wind guidance document, which is the ultimate product of SG-7.

4.2 HYPOTHESES

The hypotheses were written at a very high level to provide guidance for more detailed questioning. As presented, the general confirming or refuting of the hypotheses would provide relevant information for inclusion into the wind guidance document.

Three sets of hypotheses were developed. The first set was developed to specifically address questions related to RTA operations during descents. The second set of hypotheses, which by design is very similar to the first set, relates to questions regarding Interval Management but only in relation to a certain stage of its operation from the time of the IM assignment up to the achieve-by-point.

Other hypotheses were also developed to address questions relating to the use of potentially dissimilar forecast information, i.e., the use of forecast information by an aircraft that is either from a different forecast publication, from a different forecast source, or both.

The hypotheses are presented in their original form as developed in the noted plenary session. Modifications to these hypotheses are expected with additional input from stakeholders and from insights drawn from experiments conducted in this phase of the project.

4.2.1 RTA Related Hypotheses

Table 3 lists the four hypotheses specific to RTA performance for modern forecast models and FMS systems. Recalling that the hypotheses are written at a high level, an explanation of the intent of the wording follows.

TABLE 3

Hypotheses Formulated to Address SG-7 Questions Pertinent to RTA Operations

ID	Hypothesis
RTA-1	During NAS-wide operations, using standard publically available wind forecasts, RTA performance will be sufficient to meet ±10 seconds for 95% of RTA flights during descents to 10,000 ft MSL at Terminal Radar Approach Control (TRACON) boundary.
RTA-2	During NAS-wide operations, using standard publically available wind forecasts, RTA performance will be sufficient to meet ±10 seconds for 95% of RTA flights during descents to 2000 ft MSL at final approach fix (FAF).
RTA-3	During NAS-wide operations, the use of enhanced winds (current aircraft-sensed winds) in lieu of publically available wind forecast products prior to top of descent (TOD) will improve RTA performance by 1 sec in terms of standard deviation reduction.
RTA-4	During NAS-wide operations, the use of nine descent winds at set altitudes in comparison with four descents winds “optimally” chosen provides no significant improvement in meeting RTA requirements during descents.

The notion of National Airspace System (NAS)-wide is intended to suggest that the testing should encompass an evaluation across a diverse set of airports. The goal here is to capture the effects of different geographical conditions, route constraints, and meteorological conditions and other factors that can affect the performance of aircraft performing 4D-TBO procedures.

The notion of publically available infers the use of forecast data produced by NOAA/NCEP that are readily available to public, private, and governmental organizations without cost.

Two separate hypotheses were developed to evaluate performance at multiple altitudes because of particular interests to SG-7 and other stakeholders. The lower altitude hypothesis was of particular interest to participating National Air Traffic Controllers Association (NATCA) representatives and encouraged its acceptance. In RTA-1 and RTA-2, the reference of the ±10 seconds performance error for 95% of all flights is based on the current set of performance criteria for aircraft conducting an RTA operation in a descent [1],[6]. RTA-1 refers to operations down to altitudes of 10,000 ft and akin to track altitudes found near common TRACON merge points. The intent is to analyze flights that descend to regions of space near the TRACON boundary.

In RTA-2, the specifying of a very low altitude or down to the final approach fix is a new potential application for RTA operations. The conditions of this hypothesis have already been modified by SG-7 (changed to ~4000 ft at the initial approach fix) and are likely to change again.

Hypothesis RTA-3 presents the idea of "enhanced winds." This is an often discussed concept of utilizing locally sensed winds based on samples reported from other aircraft in the terminal area as a substitute or augmentation of forecast information. The exact definition and use case has yet to be defined, but there is an expectation of improved RTA performance if using such data.

The hypothesis RTA-4 was developed to address the questions regarding the necessity to enhance typical FMS system capabilities to meet RTA performance goals. It is expected that increasing the number of descent forecast level in the FMS would provide improved performance, but the application of additional capabilities will likely be subject to the appropriate use of those capabilities. Evaluation of different forecast sample techniques is expected as part of this consideration.

4.2.2 Dissimilar Forecast Data Hypothesis

It is not likely that all aircraft will have the same forecast information programmed in their avionics at a given time. The age, the source, and the sampling for forecast data for aircraft on identical routes may be different. A question of do all systems need to have the same forecast information has been raised by multiple parties on multiple occasions and for logical reasons.

The general named case of interest is when two aircraft are in-trail (on the same route, traveling in the same direction, and with no aircraft between them) and are assigned to arrive at a common fix at adjacent times (fixed adjacent times for RTA flights and variable adjacent times for IM flights). Could discrepancies in forecast data under these conditions lead to an unacceptable encroachment on separation minimums? See Table 4.

TABLE 4

Hypothesis Formulated to Address SG-7 Question Pertinent to Use of Common Forecast Data

ID	Hypothesis
E-1	Two in-trail aircraft conducting RTA operations with adjacent assigned times of arrival but with dissimilar wind forecasts will maintain acceptable separation without requiring external intervention.

4.2.3 IM-Related Hypotheses

Table 5 presents the hypotheses developed with SG-7 for IM operations. The terms "NAS-wide" and "publically available" should be interpreted as having the same meaning for IM hypotheses as it does for RTA hypotheses. The operational point of interest for IM hypotheses, however, is the assigned achieve-by-point (ABP). The altitude regions of interest are also the same, but there will likely be no change to the low-altitude ABP as IM operations are expected to be conducted in this domain.

TABLE 5

Hypotheses Formulated to Address SG-7 Questions Pertinent to Interval Management Operations

ID	Hypothesis
IM-1	During NAS-wide operations, using standard publically available wind forecasts, time separation performance requirements of IM operations can be met in descents to an achieve-by-point (ABP) to ~10,000 MSL at TRACON boundary in absence of forecast data specifically for the track of the traffic to follow (TTF) prior to the ABP.
IM-2	During NAS-wide operations, using standard publically available wind forecasts, time separation performance requirements of IM operations cannot be met in descents to a final approach fix (FAF) 2000 above ground level (AGL) in absence of forecast data specifically for the track of traffic to follow prior to the ABP.
IM-3	During NAS-wide operations, using standard publically available wind forecasts, the standard deviation of time separations at IM achieve-by-points is reduced by 1 sec by incorporating local/near-real-time wind data in the IM clearance in descents to 2000 AGL.
IM-4	Aircraft on routes that merge at ABP will have a greater standard deviation of time separation performance than aircraft on coincident routes in absence of forecast data specifically for track of TTF prior to the ABP.

The particular questions of interest in IM-1 and IM-2 relate to the availability of forecast information for the second participant in the operation, the traffic (aircraft) to follow.

Akin to RTA-3, IM-3 refers to the use of local or "enhanced" wind information; the definition of which is yet to be defined.

The last hypothesis, IM-4, was created to test the effects of IM aircraft flying significantly different routes and its evaluation will likely exemplify the importance of having forecast information for the traffic to follow.

There is no eluding to the level of sophistication of the IM systems to be studied to test these hypotheses. The different levels of IM system integration and IM algorithms used will have to be clearly and rationally specified as to support the formulating of the most meaningful content required in the SC-206/SG-7 guidance document.

Unlike the RTA and E hypotheses, no test matrix was developed to evaluate the IM hypotheses stated in this section. The development was not deemed necessary as SG-7 was unable to identify an assignee that was in a position to evaluate the hypotheses with new research. The test matrices for RTA and dissimilar forecast, however, are presented below.

4.3 RTA TEST MATRIX

A test matrix was developed in support of the RTCA Committee to encapsulate the range of test conditions necessary to evaluate the above RTA hypotheses. See TABLE 6. These tests examine distinct ATC, wind forecast, and equipage scenarios that differ from other tests along at least a single dimension. Each scenario can be characterized based on six independent variables: Number of Descent Forecast Points, Data Source, Forecast Sampling Method, Forecast Update Condition, Metering Fix Altitude, and RTA Assignment Distance.

TABLE 6

SG-7 Test Conditions

	Independent Variable	Values Tested	Number of Permutations
Equipage	# of Descent Forecast Points	4, 9	2
Wind Scenario	Data Source	GFS (age appropriate)	3
		HRRR 3 hour	
		"Enhanced" winds	
	Forecast Sampling Method	At planned TOD	3
		Geo-specific	
		"Optimal"	
	Forecast Update Condition	At time of RTA assignemnt	1
ATC Scenario	Metering Fix Altitude	10,000 ft, 2000 ft	2
	RTA Assignment Distance	300 NM radially from airport	1
Total Permutations			**36**
NAS-wide		Number of airports	**10**
		Samples per airport	**50**
		Total Runs	**18000**

A description of the independent variables used is provided below.

4.3.1 Number of Descent Forecast Points

Hypothesis RTA-4 examines the effectiveness of using additional forecast points relative to improved forecast altitude sample selection. It postulates that selection with four points using better wind selection techniques provides comparable or better RTA performance to using nine points selected at fixed altitudes. In order to evaluate this hypothesis, it is necessary to run tests with four levels using optimal wind selection and nine points at fixed altitudes. There are, however, a number of open questions that such a prescribed set of test cases may not answer. Does using nine points provide a statistically significant improvement versus four points when optimal wind selection is used in both cases? What level of RTA performance can be achieved using four-point fixed altitude selection? Is the type of wind forecast product a key driver in meeting RTA performance when using four or nine points? Thus, for the sake of completeness, the matrix was constructed to perform all test sets with both four and nine point wind forecast interpolation.

4.3.2 Data Sources

While Section 3 discussed the improvement in forecast accuracy that could be attained by using more recent forecasts, such improvement does not speak to the benefit of using better forecast products on RTA adherence. In order to evaluate this question, the test matrix was constructed to incorporate test scenarios for GFS forecasts as well as HRRR forecasts with look-ahead horizons of varying length. In addition, to emulate the effect of "enhanced" winds, wind samples obtained from MDCRS winds could be used as a proxy. By measuring performance when such data is used, we could then evaluate RTA performance of enhanced winds relative to a HRRR or GFS product.

4.3.3 Forecast Sampling Methods

Three methods of descent forecast sampling for RTA operations were introduced above: at planned TOD, geo-specific, and "optimal" selection. The test matrix incorporates each of these methods with varying data sources. Although none of the four hypotheses listed in Table 3 speak directly to the impact of one sampling method over another, hypothesis RTA-4 examines the impact of more wind samples vs. improved forecast sampling. To address this hypothesis, the test matrix was configured to evaluate both "optimal" and geo-specific wind sampling at fixed altitudes. To replicate the procedure of some airlines, an additional test configuration was added to allow for sampling points at fixed altitudes under the top of descent (TOD) location.

4.3.4 Forecast Update Conditions

There are a number of potential configurations that could be used to model impact of forecast updates including prior to take-off, periodically, at TOD or at a predefined radial distance. In order to limit the dimensionality of the test matrix, the forecast update location was specified to occur at a distance of 60 NM from the location of RTA assignment.

4.3.5 Metering Fix Altitudes

Hypotheses RTA-1 and RTA-2 both directly address the ability of FMS equipped aircraft to adhere to their controlled time of arrival (CTA) assignments down to specified altitudes. To investigate each statement, the metering fix altitudes in the test matrix were set to altitudes near 10,000 ft MSL and 2000 ft AGL. Each STAR under test was examined in order to identify appropriate waypoints to mid-altitude and low-altitude metering fixes nearest to 10,000 and 2000 ft, respectively.

4.3.6 RTA Assignment Distance

Based on the recommendations of NATCA representatives, a radial RTA assignment distance of 300 NM away from the destination airport was used for all test cases.

4.3.7 Number of Trials

To complete the test matrix, 10 major U.S. airports of interest will be identified for study. These airports will be selected for geographical diversity and potential for providing an ample number of test cases and interesting route configurations. Since the weather conditions will also vary on a daily, if not hourly, basis, it is important to select a sufficient number of samples to draw statistically meaningful conclusions. The real-time nature of the planned simulations imposes an upper bound on the number of simulations that would be feasible. Considering both these factors, a decision to use a sample size of 50 flights per airport was chosen.

4.4 DISSIMILAR FORECAST TEST MATRIX

While the above test matrix examines RTA performance under a number of scenarios to address the RTA-related hypotheses, it cannot be used to evaluate hypothesis E-1 regarding aircraft separation between two in-trail aircraft. As such, an additional test matrix was created to address this issue. Given the number potential confounding variable associated with using two different forecast products or a common product with different look-ahead horizons, we chose to use a controlled difference in wind forecast error instead. The matrix outlines a concise set of test conditions where the number of forecast levels, the forecast data source, the sampling method, and the update condition are all held constant while the wind forecast error and distance traveled under an RTA operation at cruise are manipulated. The in-trail aircraft separation test matrix is shown in Table 7.

TABLE 7

In-Trail Test Matrix

	Independent Variable	Values Tested	Number of Permutations
Equipage	# of Wind Forecast Points	4	1
Wind Scenario	Forecast Data Source	HRRR 2	1
	Forecast Sampling Method	Geo-specific	1
	Forecast Update Condition	At time of RTA assignment	1
	Difference in Forecast Values	0 kt, 10 kt, 20 kt, 30 kt, 40 kt	5
ATC Scenario	Metering Fix Location	TRACON Boundary	1
	RTA Assignment Distance	150 NM, 250 NM, 350 NM	3
Total Permutations			**15**
		Number of airports	**10**
		Samples per airport	**5**
		Total Runs	**750**

4.4.1 Difference in Forecast Values

Runs will be repeated using the same forecasts, but the direction and magnitude at selected descent forecast levels is modified to control different along-track head wind forecast errors. The added errors will exceed at least three standard deviations in forecast errors as identified earlier in this work.

4.4.2 RTA Assignment Distance

To measure the effect of the integrated errors, the starting location of the procedure is placed at distances 100 NM apart. This variable is modified with the expectation of seeing increased encroachment with greater assignment lengths.

The elements of Section 4 were developed as the outcome of the support provided to RTCA SC-206/SG-7. The hypotheses and test matrices were created to define areas of research to be conducted by parties that participate in SC-206/SG-7 activities. It was understood by SG-7 that not all test conditions or proof of all hypotheses would be evaluated by one organization. Lincoln Laboratory was requested to

evaluate questions related to RTA performance as motivated by the SG-7 test matrices, results of which are discussed in the next section. Investigations of the hypotheses related to the IM and in-trail operations were left unassigned by SG-7 as there was no party available that was both willing and having the capacity of resources at that time to perform the required research. There is still strong interest by SG-7 to have the complete suite of research conducted in the future.

5. RTA PERFORMANCE AS A FUNCTION OF FMS CAPABILITY

5.1 OVERVIEW

In consideration of the research goals established in supporting RTCA as directed by the sponsor and to address the particular question covered under Task 3 of the program, an analysis was conducted to evaluate the performance impact of an FMS that can utilize nine levels of descent forecasts as opposed to one that can only utilize four levels of descent forecasts. The analysis was motivated by the contents of the test matrix developed with SC-206/SG-7 described in the previous section.

As described in Section 2.2.2 *Enhancements to FMS Capabilities*, the modified FMS provided the capability of programming up to nine descent forecast levels (DFLs), which was required in order to conduct this analysis.

5.2 HYPOTHESES

The tests described herein are designed to specifically analyze the impacts to RTA performance of nine versus four DFLs using the hypotheses in Table 8 below.

TABLE 8

Hypotheses Composed to Evaluate Number Effect of Available Descent Winds

ID	Hypothesis
DL-1	The RTA time error bias is at least 1 second closer to 0 for flights that use nine optimally selected descent winds as compared to flights that use four optimally selected descent winds
DL-2	The standard deviation of RTA time error is at least 1 second smaller for flights that use nine optimally selected descent winds as compared to flights that use four optimally selected descent winds
DL-3	The RTA time error bias is closer to 0 for flights using forecast winds selected to minimize the estimated headwind error versus winds selected to minimize estimated magnitude error
DL-4	More than 90% of individual flights that use nine optimally selected descent winds as compared to the same flight that uses four optimally selected descent winds have a smaller absolute RTA time error

5.3 TEST MATRIX

The test matrix developed for this analysis is very similar to the one developed for SC-206/SG-7 activities described in the previous section. It is principally a subset of that matrix but is specific to address the hypotheses presented in this section and with a slightly modified set of operating conditions. See Table 9. It also considers the baseline performance cases where either no forecast information data or perfect forecast information (truth data) is provided. The design, creation, and analyses from these tests will lead to greater refinement of those tests required for SC-206/SG-7 work.

TABLE 9

FMS Capability Evaluation Test Matrix

	Independent Variable	Values Tested	Number of Permutations
Equipage	# of Descent Forecast Points	0, 4, 9	3
Wind Scenario	Forecast Data Source	HRRR 3 hour	1
	Forecast Sampling Method	Optimized for HW Optimized for Magn Use truth data	3
	Forecast Update Condition	340 NM from dest	1
ATC Scenario	Metering Fix Altitude	10–15 kft	1
	RTA Assignment Distance	280 NM from dest	1
Total Permutations			**9**
NAS-wide		Number of Airports	**10**
		Samples per Airport	**50**
		Total Flights	**4500**

5.4 TEST DESIGN

Enhancements to the existing simulation framework were performed as described in Section 2.2 in order to support the experiments outlined in the test matrix. As stated earlier, all scenarios experience atmospheric conditions as measured during flight.

After fusing, correlating, and evaluating recorded flight information as described in Section 2.2.3, the next step is scenario creation. This process exercises a number of steps with the ultimate goal of producing all the required scenario configuration files used to conduct each individual experiment. A flow diagram of the process is shown in Figure 25 and a description of each component follows.

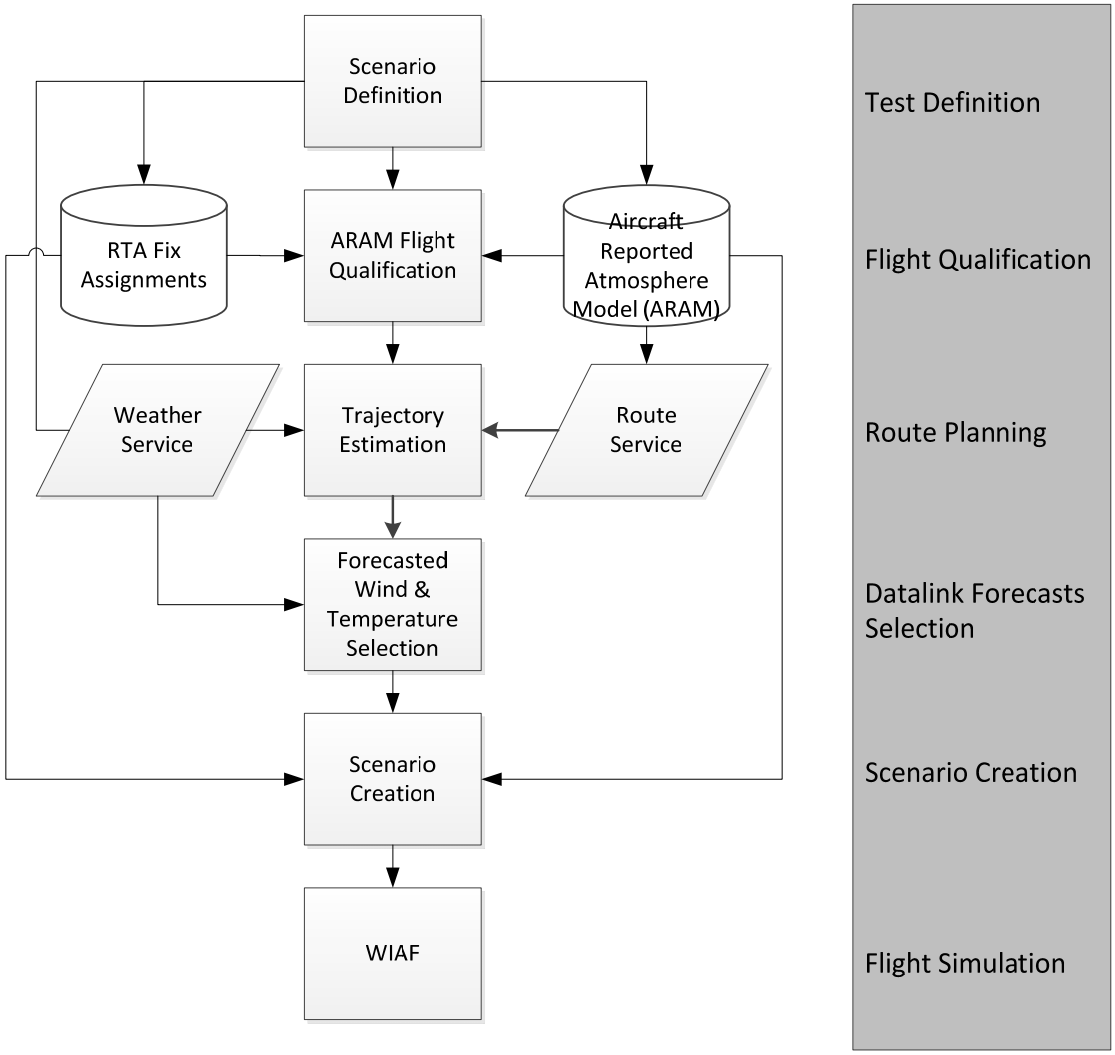

Figure 25. Scenario creation flow diagram.

5.4.1 Scenario Definition

The scenario definition is a collection of parameters required for each experiment. Elements of these common parameters and those that are specific to an experiment are used to identify candidate ARAM flights as well as establish characteristics of agents used in the experiment. Some common experiment parameters include the specific aircraft type being modeled, its current weight, FMS type, specifics of the programmed route, what forecast information has been uploaded to the FMS and the source of the forecast, etc. An example of the scenario definition profile is provided in Table 10.

5.4.2 Flight Qualification

Not all the ARAM flights stored in MAFID can be used in these experiments. Recall that an ARAM flight is a contiguous segment of a particular flight that both remained on its assigned route and had a complete set of MDCRS data sampled over the same period. There may be more than one ARAM flight (or flight segment) created from a given flight if the aircraft temporarily deviated from its route or there were periods of missing MDRCS data.

The parameters in the test matrix such as RTA assignment distance and others are used to filter the available ARAM flights such that their characteristics satisfy the minimum set of conditions needed for an experiment. For example, if the starting location of the simulated aircraft is on the assigned route 340 NM radially from the destination, an ARAM candidate flight must have the start of its data set begin earlier in the route than this point. That same candidate flight must also have is ending point occur after the RTA fix location on that route (plus some additional buffer distance). The particular RTA fix used the filtering process and applied in the experiment is route dependent. The fix is selected by examining each STAR and identifying the fix where the expected crossing altitude was between 10,000 and 15,000 ft. Additional qualifiers such as continuous cruise altitudes (±1000 ft) prior to the estimated TOD point from the track data must also be met.

5.4.3 Trajectory Estimation

The modelled RTA operation used in these simulations presumed that at some distance prior to arriving at an expected sequence and scheduling freeze horizon, the flight crew would have requested a forecast update from their respective airline operations center (AOC). As part of this request, the onboard automation would have downlinked to the AOC the current cruise altitude along with an estimate on the remaining trajectory. It was presumed that the AOC would likely use this estimated trajectory to perform its forecast selection.

Not knowing what the theoretical estimated trajectory downlinked to the AOC is or if a particular AOC would use some other means to estimate the remaining trajectory, we used a subdivided set of track reports from the TFMS data as the estimate for the remaining trajectory.

5.4.4 Forecast Selection

One of the basic assumptions employed in these experiments was that the modelled AOC did not have a forecast selection process that considered if the aircraft requesting a forecast update was participating in an RTA operation. That is, the simulated AOC, when selecting the forecast to provide to its aircraft, did not limit the range of descent forecast altitude levels to consider, which could be done if one considered the estimated RTA fix crossing altitude. Thus, the simulated forecast selection process considered the full range of altitudes from cruise to destination altitude when analyzing which forecast data to deliver to the aircraft. We believe this is a reasonable assumption and consistent with current AOC operations.

Consistent with the test matrix, we provided either no forecast data or forecast data for both cruise and descent. Under all cases, if forecast data was programmed into the FMS, it included the expected wind magnitude, wind direction, and ambient air temperature for each entry.

Forecast Age

The notion of forecast age is a concern when using model data when sampling a forecast. We define the age of the forecast as the difference in wall clock time from when a forecast is requested minus the named model cycle run time for the forecast being used. In this work, which uses the HRRR model, forecasts are not available from the NOAA distribution site until approximately 1 hour from the named cycle run time due to generation and delivery delays. We add transport and processing delay of 2 hours to model AOC operations, which means the youngest age of a forecast that can be referenced is 3 hours old. The oldest a forecast could be is 4 hours since HRRR forecast sets are generated every hour.

For requests made to the weather forecast data service, the simulated wall clock time of the request minus three hours is passed as a parameter that limits the latest model cycle runtime forecast set that should be used as the source when providing the forecast data.

Cruise Forecasts

The process for selecting forecast data for cruise waypoints is straightforward. When HRRR forecast data is the source, the estimated trajectory is used to determine where the aircraft is expected to cross its cruise waypoints in time and space, and a request for conditions at these locations are made to the weather forecast data service. When MDCRS as truth is used as the forecast source, the conditions at each of the cruise waypoints are interpolated from the MDCRS data. Forecast for cruise waypoints are only provided at the cruise altitude and for waypoints along the route from the starting location up to two waypoints beyond the estimated TOD location.

Descent Forecasts

There are a number of techniques used in the airline industry today to decide what descent forecast information should be provided to an aircraft's FMS. Three commonly used approaches are to select

forecast values geographically situated over the estimated TOD, geographically situated over the destination, or geographically along the descent profile. These approaches typically make their selections at fixed, company-specific altitudes, regardless of the location selection technique.

We decided that in order to conduct a fair comparison on the efficacy of the use of four versus nine DFLs, it would be prudent to perform the forecast selections along the predicted nominal descent trajectories and not over fixed geographical points. Additionally, because the atmospheric conditions most relevant to correctly predicting a descent trajectory do not occur at the same altitudes with each flight, we decided to permit the DFL altitudes, and thus locations, to vary on each flight chosen in an optimal manner described below.

If providing descent forecasts, the selection process used the same general technique regardless of the number of descent forecasts to use. That is, we employed an optimization approach using a decimation (or "greedy") algorithm, which minimized some cost function, which resulted in a set of forecasts chosen to provide to the FMS. In addition to testing RTA performance with four and nine DFLs, we also analyzed RTA performance on each of these cases using two different cost function profiles in the optimization process. One cost function tried to optimize for the most accurate wind magnitude profile expected along the descent. The second optimized for the most accurate representation of the along-track headwinds expected along the descent.

When using HRRR as the source for descent forecasts, the sampled forecasts are collected in the same manner as used for cruise waypoints. However, when using MDCRS data as the source, the closest meteorological samples in terms of altitude are used as the forecast samples.

Optimization Process

The decimation ("greedy") optimization algorithm is a simple and fast approach to find solutions to large set problems and is used in many disciplines. For example, it is similar to that used by Ahmed et al. in their descent forecast selection [11]. This algorithm, as with many optimization techniques, does not guarantee it will identify a globally optimal solution for all problem types, and this is consistently true for the cases studied here. However, analysis of its application in this work has shown that it does appear to find a near global optimal solution and is quite suitable for this effort.

As applied to this work, a set of points define a desired profile to replicate is created from sampling time-associated wind forecast data at 1000 foot intervals along a predicted descent trajectory. The highest altitude in the set was the cruise altitude for a given flight and does not vary during the optimization. The lowest altitude was the first altitude rounded to 1000 feet that exists above the destination airport.

The cost function was the sum of the absolute error between the desired profile and the interpolation at the desired profile altitudes using the reduced-point set. The reduced-point set began as the desired profile set but was reduced by one point in each major iteration. The point selected for removal was that point whose removal amongst all the remaining points produced the smallest numerical

cost if removed. These iterations continued until the number of points remaining equaled the desired number of DFLs (four or nine in our case).

The optimization progression is demonstrated in Figure 26. In this example, the optimization was attempting to fit to the calculated headwinds that were estimated on that descent. These winds range from an estimated tailwind of 60 kts to a peak headwind of around 20 kts. Starting with the full complement of available DFLs selectable for that particular flight, we can see the effect on the fit as the points are removed until reduced to the desired number of DFLs. The point at the cruise altitude is never removed as it is always available as part of the cruise forecast. Note that, in this particular example, the lowest altitude retained in the reductions to both nine and four DFLs happens to be the lowest altitude from the full set. This is not always the case.

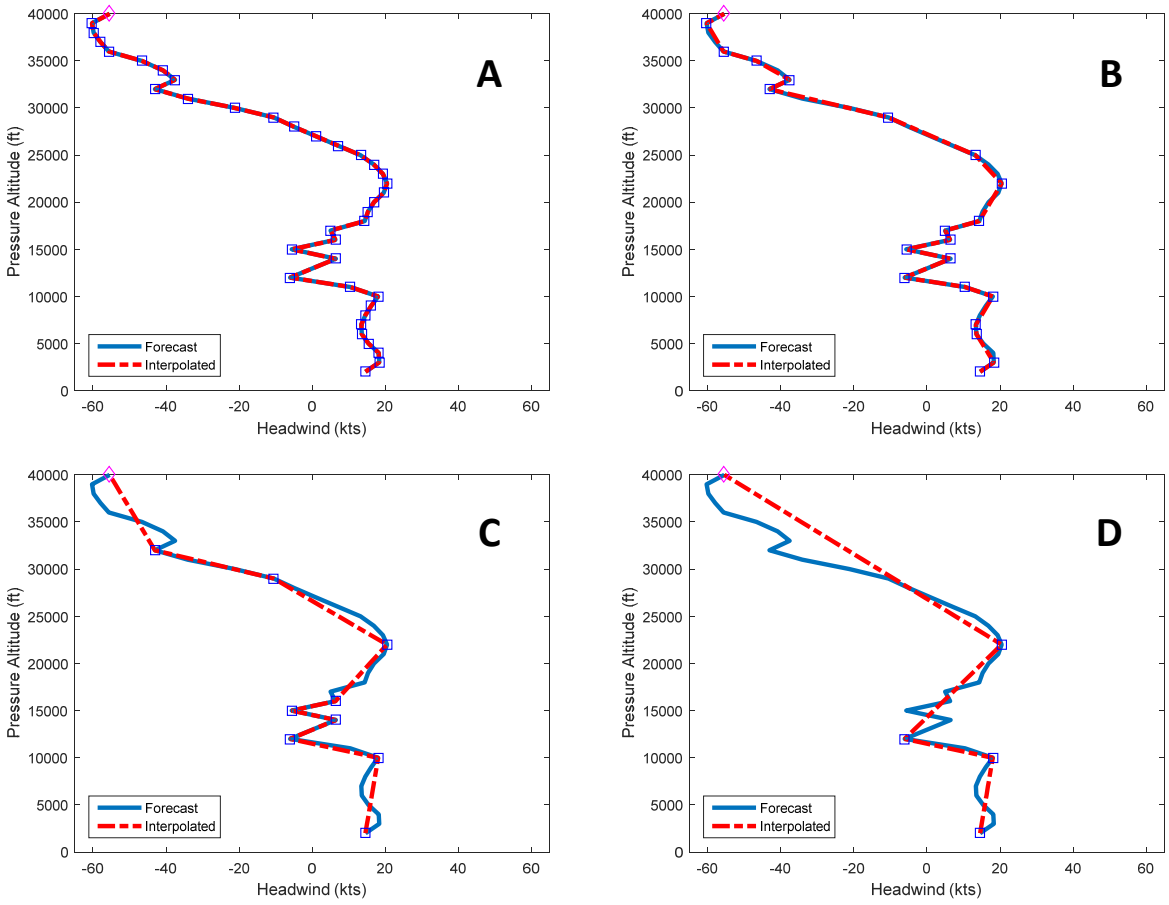

Figure 26. Progressive example of forecast selection decimation process for a flight into KPHX optimizing for the best fit to the headwinds expected along the descent trajectory. Plot (A) shows the desired profile (blue) and the 37 starting altitudes, plot (B) the fitted profile made up of 20 points, plot (C) the same profile reduced to nine points, and plot (D) the fitted profile made up of four descent forecasts.

Using the same flight as presented in Figure 26, we can view the resulting profile when attempting to fit the DFLs to the magnitude of the wind forecasted at each altitude. The plots in Figure 27 demonstrate the fits using nine DFLs (A) and four DFLs (B). The forecasted peak magnitude for that flight occurred just below cruise altitude with a coincidentally relatively linear change in wind magnitude throughout the descent.

It is interesting to note that the three highest DFLs chosen in the two 4 DFL cases presented are significantly different, depending on which cost function was utilized. From the solution presented in Figure 26(D), we see that the forecasted atmosphere conditions were selected at FL220, 12,000 ft, 10,000 ft, and 2000 ft at their respected geographic locations, but for the solution shown in Figure 27(B), values at FL320, FL290, FL220, and 2000 ft were selected.

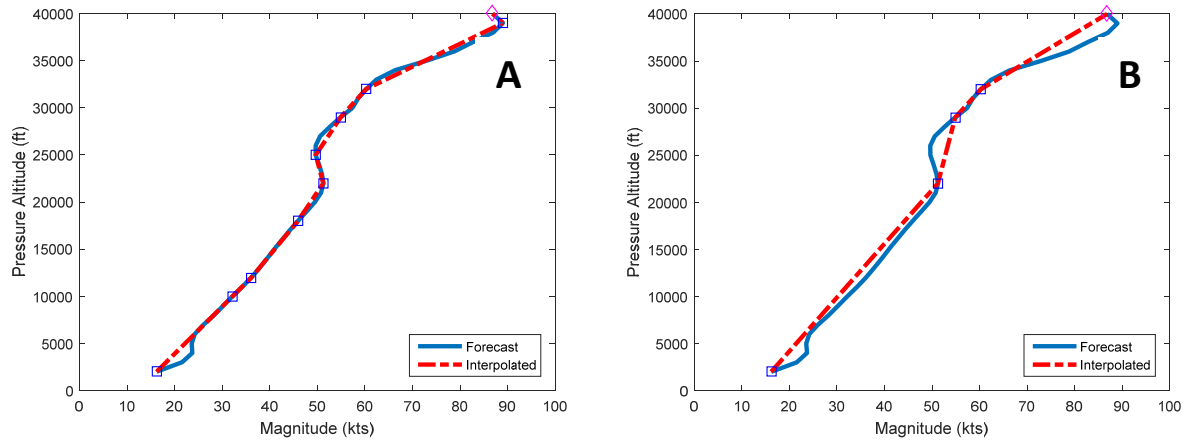

Figure 27. Interpolated fit to forecasted "truth" after solving for nine (A) and four (B) DFLs for the same flight presented in Figure 26 but with wind magnitude used in the cost function instead of headwind magnitude.

Each remaining trajectory was partitioned into a cruise phase and a descent phase of flight. The MDCRS wind and temperature data associated with each phase was collected to be used later as a basis of truth for our simulation.

5.4.5 Scenario Creation

The MDCRS track data was truncated to a point just beyond the initial starting location of the simulation. To ensure a complete set of reference data over the airspace of interest, MDCRS cruise weather samples from the descent portion of the track extending two waypoints beyond the estimated TOD were added on the cruise reference weather data. This ensured that reference weather data was present even if the aircraft descended later than what was observed in the actual track. Once the simulated aircraft began to descend, the weather data sampled during the MDCRS aircraft descent was applied. The

metering fixes were identified by examining each STAR and selecting a fix where the expected crossing altitude was between 10,000 and 15,000 ft. These points were hand-selected and mapped to a database table that stored the selection of every potential approach. The metering fix of interest was retrieved from the database based on the approach procedure under test. The RTA assignment point was set to correspond with a radial distance of 280 NM from the destination airport. The initial position of the flight was defined by identifying the point along the route that was 60 NM prior to the location of RTA assignment at a radial distance of 340 NM from the destination airport.

Once these points were defined, a complete scenario profile was written into an XML file that defines a portion of the scenario for our agent-based real-time simulation. This file contains the identified wind samples from the cruise and descent phases of flight, reference data for the cruise and descent phases of flight generated from the MDCRS records, the route to be flown by the aircraft, and the to be assigned RTA fix location. There are a number of additional parameters that define the specifics of the simulation scenario, including the type of aircraft, its weight, the amount of fuel initially on board, the initial speed of the aircraft, and the cost index assumed by the FMS. An additional file is created to provide supplementary aircraft, wind, and trajectory information for each scenario. The details of this scenario are provided in Section 5.6.

5.5 TEST DESCRIPTION

A set of test scenarios was developed to provide an assessment of the research objectives. Each scenario can be characterized based on six parameters, including the number of descent forecast levels, data source, forecast sampling method, forecast update condition, metering fix altitude, and RTA assignment distance. The number of wind forecast levels used to calculate the wind interpolation for the descent phase of flight was limited to either four or nine points. In order to assess the potential improvement associated with using standard FMS wind profile adjustments, we also developed a set of test cases to evaluate RTA performance under zero wind conditions.

The HRRR wind forecast models were interpolated both spatially and temporally to supply a wind forecast at the appropriate four-dimensional locations. With the interpolation in place, it was necessary to identify the locations where the winds will be sampled. This location was defined by the trajectory of the TFMS track. The wind selection algorithm described in the previous section was then used to collect the required number of samples. Cruise winds were updated at every waypoint along the en route portion of the trajectory at the expected crossing times of each fix. An RTA assignment distance of 280 NM away from the destination airport was chosen based on the recommendations of NATCA representatives. The descent wind forecast was also issued at the time of RTA assignment. A final approach point was chosen for all possible approach procedures to select a metering fix at or near 10,000 ft MSL. To complete our tests, we identified 13 major U.S. airports of interest. In all cases, the number of tests that could be performed at each airport was limited by the number of qualified MDCRS tracks available from our data sources for each airport.

TABLE 10

Scenario Definition Parameters

Simulation Component	Parameter	Values Tested
Aircraft	Aircraft Type	B757-200
	Gross Weight	17,0000 lbs
	Zero Fuel Weight	15,0000 lbs
	Cost Index	50
	RTA Tolerance	6 seconds
Forecast	Age	>3 hours since publication truth data
	Update Condition	60 NM prior to RTA assignment location
Trajectory	RTA Assignment Distance	280 NM radially from destination
	Acceptable RTA Fix Altitudes	10–15 Kft MSL
	Initial Aircraft Starting Point	60 NM prior to RTA assignment location

Two scenario files were used to define the additional aircraft, wind, and trajectory information needed to perform the simulation. An XML file was used to define the trajectory, forecast, and wind truth data needed for simulation. An additional file was used to provide a set of reference parameters necessary to define the characteristics of the aircraft and systems under test. A summary of the parameters used is shown in Table 10. An RTA assignment time selection criterion relative to the RTA window estimated by the FMS was defined in order to perform the simulation. This criterion was defined to be 25% of the difference between the latest and earliest possible arrival times plus the earliest possible arrival time from the FMS RTA window estimate.

While the TFMS data source provides a rich pool of potential flights, many of them were unsuitable for testing. Thus, the data pool was associated and filtered to ensure that all simulated routes met our test design standards described in Section 2. An initial association process matching TFMS, ASDE-X, and MDCRS records was performed. Flights with a common set of records were processed through the route analyzer process described earlier. A set of scenario parameters was used to select the qualified routes that were applicable to the test case of interest. These parameters included flights between February 1, 2016 and March 31, 2016 whose tracks originated beyond the 280 NM radial boundary of the airport of interest. The parametric conditions also specified that the tracks extended 10 NM beyond the final RTA meter fix. This distance of 280 NM was chosen instead of 300 NM to reduce the runtime of the simulation while still providing conditions that were similar enough to NATCA representatives' recommendations to provide conclusions of interest.

Additional filtering was performed on each flight to disqualify those whose cruise segment varied in altitude by more than ±1000 ft. In some cases, an RTA metering fix could not be identified due to insufficient information from the route. When this occurred, the flights were also disqualified from the test pool. As a result of the process, the number of flights available for testing shrunk considerably. The number of routes available at each stage of the process for each document is documented in Table 11.

TABLE 11

Filtering Counts Leading to Final Set of Simulated Flights

Airport	TFMS Flights	Associated Flights w/ MDCRS	Route Analyzed Records	Criteria Qualifying Flights	Flights Simulated
KMEM	17372	4489	3771	147	0
KMDW	17773	2478	2243	77	1
KDEN	48213	2248	1907	99	68
KPHX	38528	1744	1450	73	27
KATL	77033	3541	1248	56	17
KBOS	29571	679	490	35	24
KEWR	35610	529	485	36	0
KORD	73012	346	178	1	0
KIAH	41138	229	128	0	0
KCLT	47040	467	116	1	0
KDFW	60691	221	105	0	0
KJFK	38965	912	65	0	0
KSDF	12058	72	46	0	0
Total	**537004**	**17955**	**12232**	**525**	**137**

5.6 TEST RESULTS AND ANALYSIS

Samples of results for simulated flights into KDEN are shown in Figure 28. In these particular samples, no forecast information was provided to the FMS. From the figure (top left pane), we see a large variety of headwind conditions experienced amongst the flights. This is expected as there is an ample

distribution of eastbound and westbound arriving flights. The assigned cruise levels for these flights ranged from FL280 to FL410.

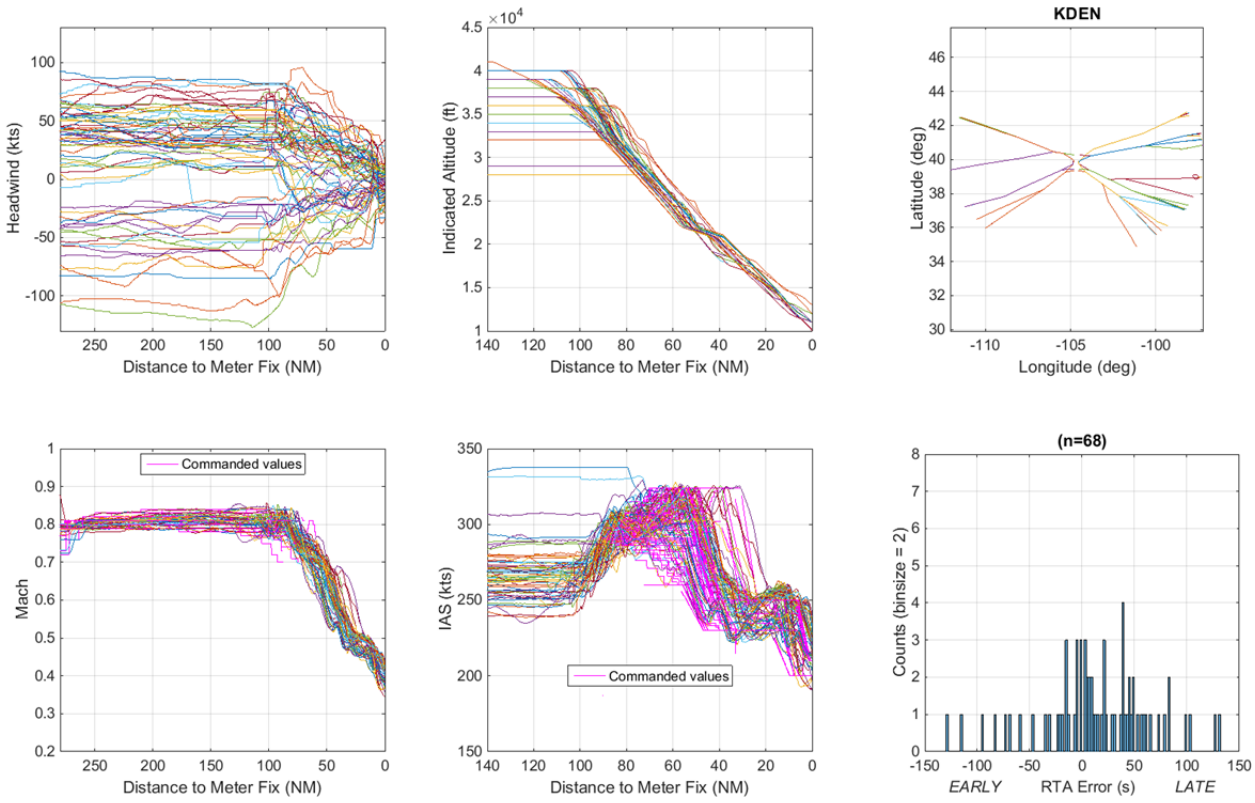

Figure 28. Simulation of flights into KDEN without forecast information provided.

Given that sensed winds and temperature are the only conditions known about the atmosphere except for the barometric correction for the destination airport, it is not surprising to see a commonality amongst the vertical descent profiles in the top middle panel given the cost index among all flights was assumed to be the same. As seen in the histogram at the bottom right of the figure, we see a wide distribution in the arrival time errors, which is expected considering the lack of forecast information.

The results of flights flying under the same atmospheric conditions but with cruise forecasts and four descent forecast levels (DFL) data provided are presented in Figure 29. The DFLs were selected based on an optimized fit to wind magnitude. In contrast to the previous figure, a range of top-of-descent locations is seen as expected given the varied descent wind conditions. Also clearly evident are visibly reduced variances in the target Mach and indicated air speed (IAS) targets provided by the FMS. This is

an indicator of correct prediction and managing of future conditions. As indicated in the histogram in the bottom right panel, all but three of the 68 simulated flights arrived at their RTA fix within 10 seconds of their assigned RTA time, i.e., far superior performance to the previous case.

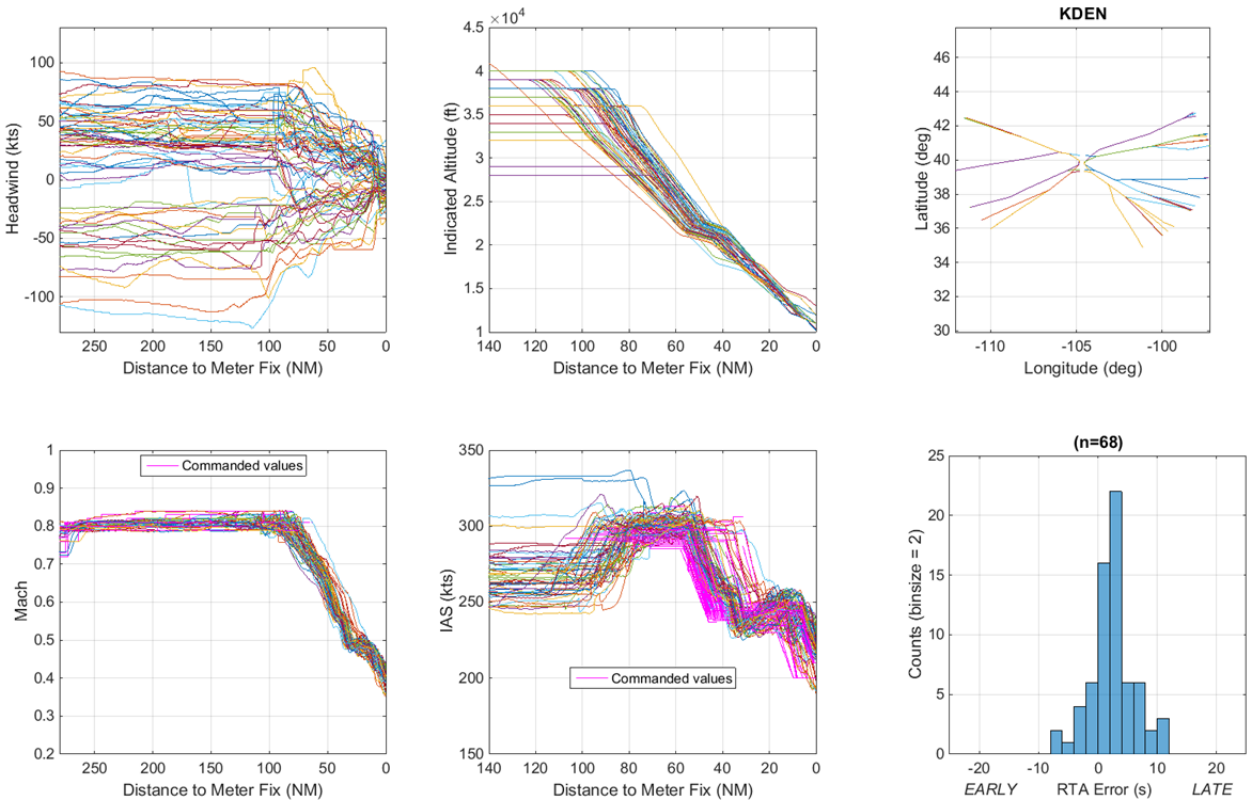

Figure 29. Simulation of flights into KDEN with cruise and four levels of descent winds.

Similar behaviors and performance can be seen in the plots of Figure 30, which demonstrates results for the same flights but with nine levels of descent forecast data provided to the FMS.

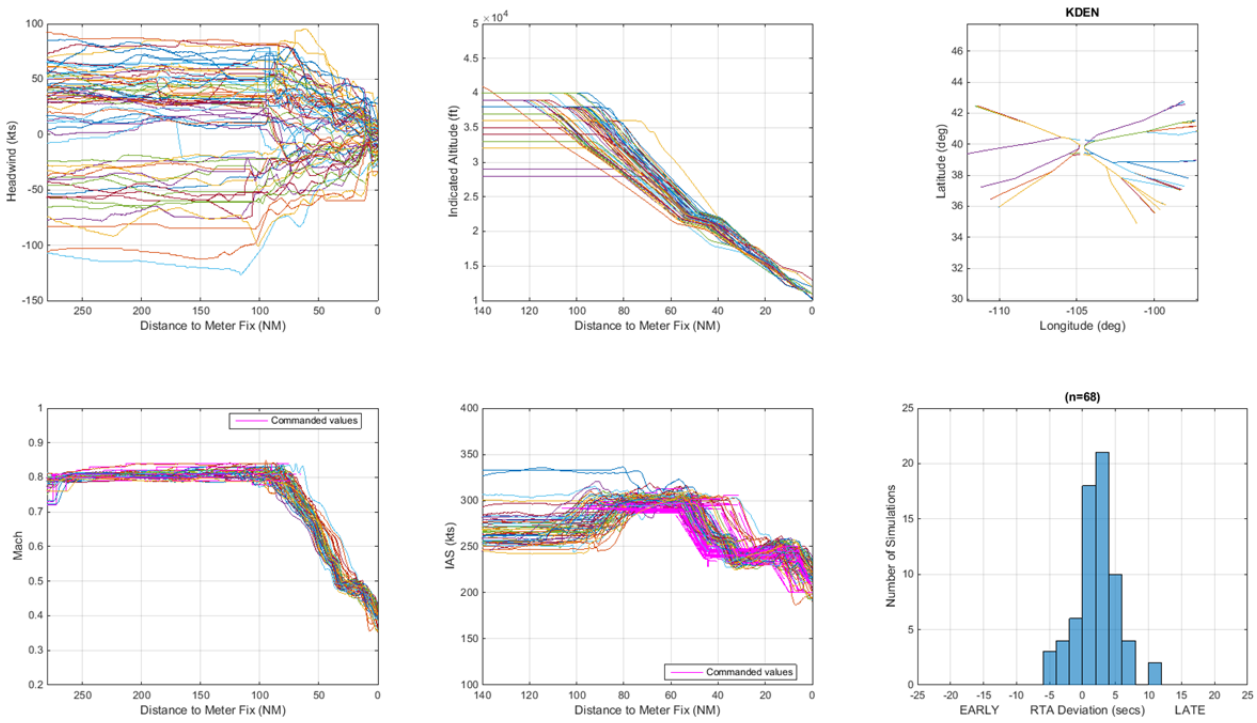

Figure 30. Simulation of flights into KDEN with cruise and nine levels of descent winds.

5.6.1 Adherence to Assigned Times of Arrival

As described in the test description, the principal design parameters that were modified were the number of descent forecast levels, 9, 4, or 0 (0 indicates neither cruise nor descent forecasts), and the wind selection optimization technique fit to headwind or magnitude.

Analysis for the aggregated sets across all destinations is provided below. Beginning with the baseline no forecast case, we present its histogram of RTA arrival time errors in Figure 31. The shaded areas around 0 seconds in this histogram (as in all of the histograms below) indicate the ±10 second arrival error region as a reference. We use this reference as it is part of the desired performance criteria proposed for RTA descent operations.

Without forecast information, we observe that approximately 25% percent of the flights reproduced arrived within the ±10 second window; 95% of these flight arrived between –101 seconds early and 112 seconds late of their assigned RTA times.

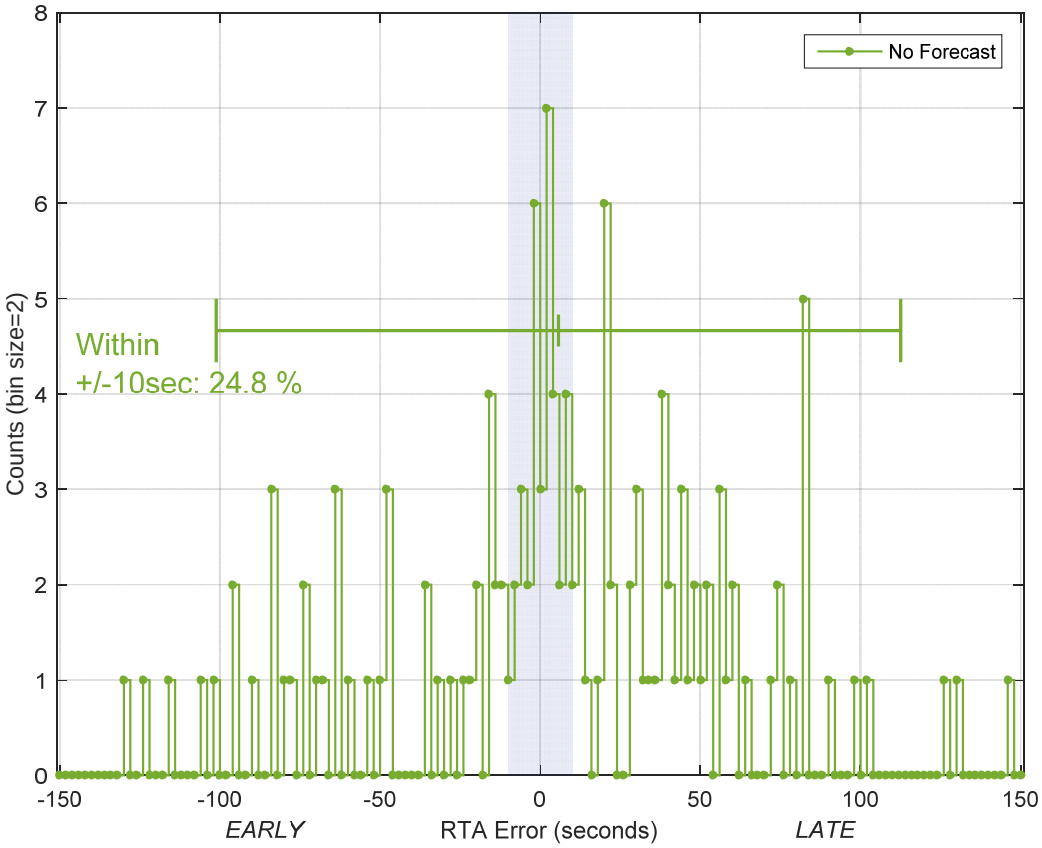

Figure 31. Histogram of RTA errors for 137 flights across all analysis airports when no forecast information is provided. The whisker bars show 2std around the mean.

Results for time arrival errors for four and nine DFL conditions when the forecast selection optimization for minimizing headwind error is chosen are presented in Figure 32. It also shows the RTA errors of the best possible performance using this selection technique by using nine DFLs and utilizing truth data in place of forecast.

As seen in previous simulations reported in earlier phases of this work, and as reported in the 2011 Seattle 4D-TBO flight trials [12], there is clearly a bias in the aggregate for the flights to arrive late relative to the assigned RTA time. In the flight trial data, the bias was approximately 9 seconds. The system under test in those trials was the B737 aircraft using General Electric (GE) FMSs with a coarser and less accurate forecasting model based on the Rapid Update Cycle (RUC) model, which is a predecessor to RAP and HRRR. In our experimental results, the bias for this four DFL case is 3.7 seconds

and for the nine DFL case, it is 3.2 seconds. Table 12 lists the basic statistics of the aggregated set of simulated flights for each of the experimental cases.

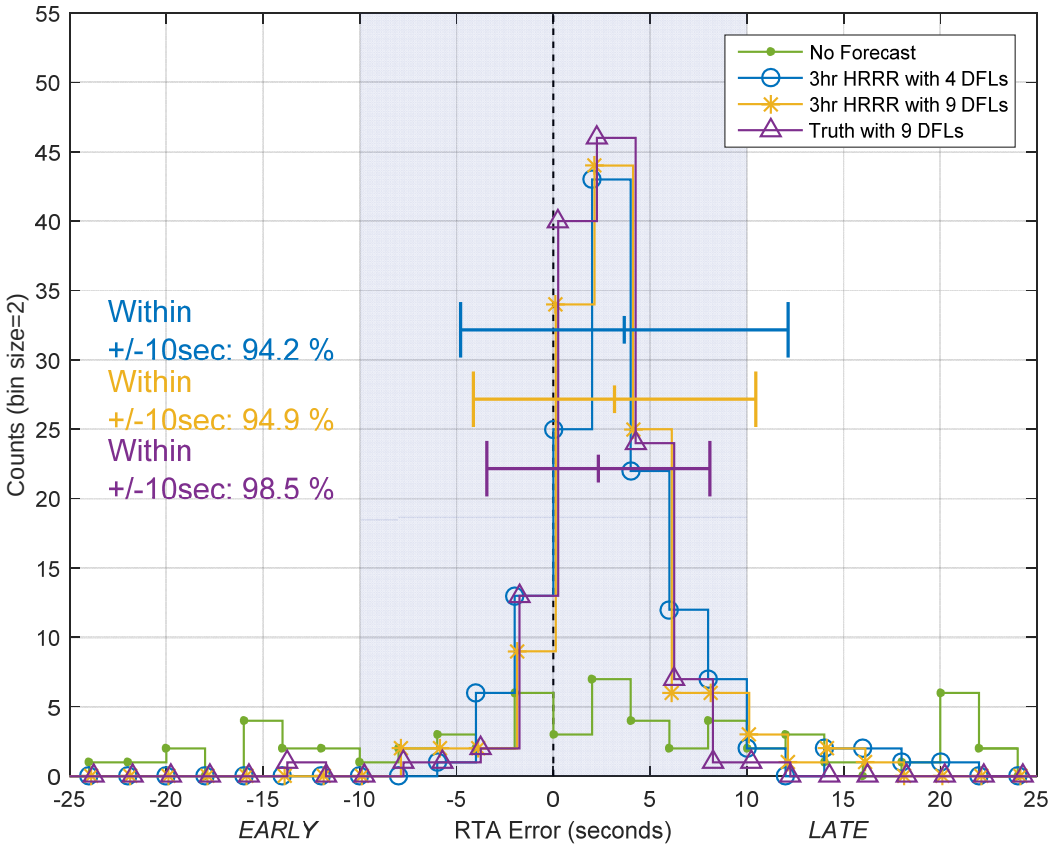

Figure 32. Histogram of RTA errors with the DFL selection technique set to Optimize to Headwind Magnitude profile. Whisker bars show 2std around the mean. (Blue) Selected from 3-hour forecast using four DFLs. (Yellow) Selected from 3-hour forecast using nine DFLs. (Purple) Selected from truth data using nine DFLs. (Green) No forecast information provided.

Figure 33 shows the histograms for the four and nine DFL cases where the forecast selection optimization to minimize wind magnitude error was chosen. Surprisingly, this choice of optimization produces biases that were smaller than those produced from the headwind optimized cases. The biases under these conditions were 3.1 seconds for both the four and nine DFL cases using forecast data. Only slight improvement in performance was observed in the nine DFL truth data case compared to the nine DFL forecast case.

As seen in Table 12, when comparing optimization to profile headwind versus profile wind magnitude, there is a minor reduction in the standard deviation for the four DFL cases (from 4.2 seconds to 4.0 seconds), but there is a slight negative effect on the standard deviation for the nine DFL case, increasing it from 3.6 to 3.7 seconds.

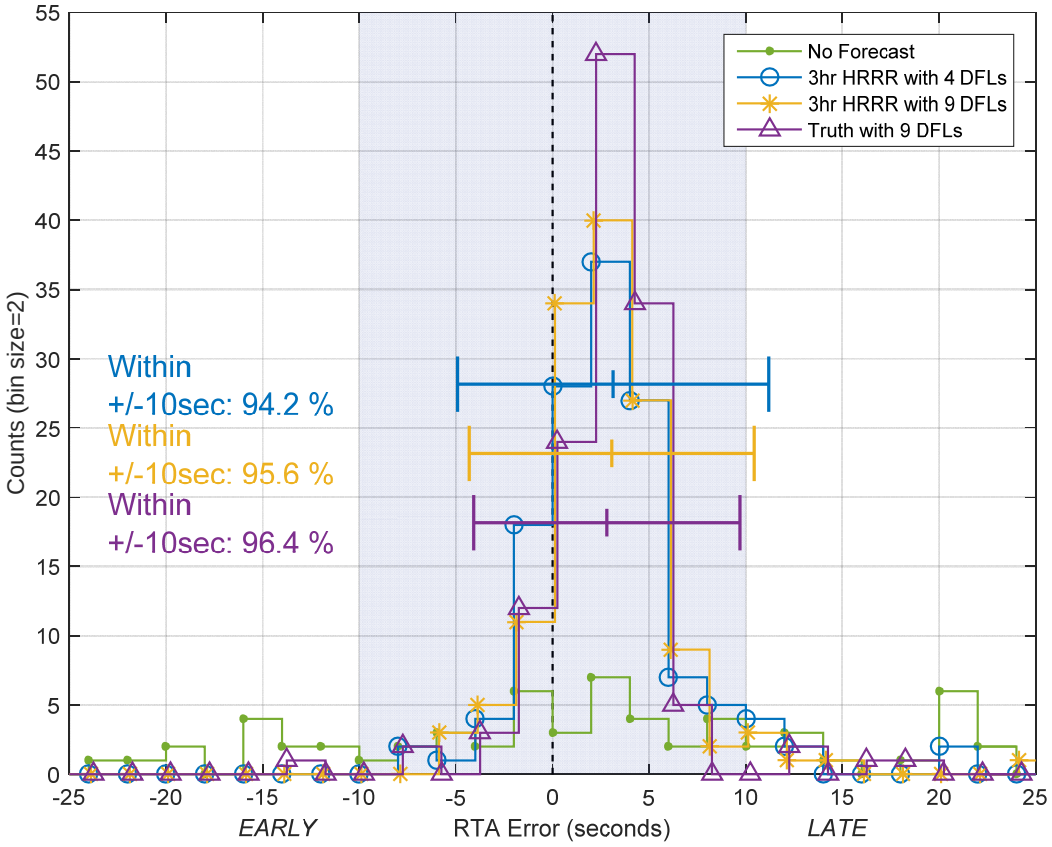

Figure 33. Histogram of RTA errors with the DFL selection technique set to Optimize to Wind Magnitude profile. Whisker bars show 2std around the mean. (Blue) Selected from 3-hour forecast using four DFLs. (Yellow) Selected from 3-hour forecast using nine DFLs. (Purple) Selected from truth data using nine DFLs. (Green) No forecast information provided.

TABLE 12

Overall RTA Error Statistics as Function of Forecast Information Source, Selection Technique, and Number of DFLs

RTA Error Mean (std) [seconds]

Selection Technique	No Forecast	3-Hour HRRR Forecast # Descent Forecast Levels 4	3-Hour HRRR Forecast # Descent Forecast Levels 9	MDCRS as Forecast # Descent Forecast Levels 9
Optimized Headwind	5.7 (53.4)	3.7 (4.2)	3.2 (3.6)	2.3 (2.9)
Optimized Magnitude	5.7 (53.4)	3.1 (4.0)	3.1 (3.7)	2.8 (3.4)

Under the given set of conditions tested, both the biases and standard deviations were reduced under all test conditions when using nine DFLs instead of four DFLs. However, these reductions were not significant, and we can establish that none of the first three first hypotheses (DL-1, DL-2, and DL-3) were substantiated.

5.6.2 Flight-by-Flight Performance Comparison

To perform a fair evaluation of the results to validate hypothesis DL-4, we inspect the results in two ways. Recall that in DL-4, the expectation is that on any given flight the use of nine DFLs should produce a smaller RTA time error (RTA TE) compared to four DFLs most of the time (i.e., >90% of the time).

The first way we shall look at this is with a generalized comparison of a count of the number of cases where $|\text{RTA TE}|_{9\ \text{DFL}} < |\text{RTA TE}|_{4\ \text{DFL}}$ versus $|\text{RTA TE}|_{9\ \text{DFL}} \geq |\text{RTA TE}|_{4\ \text{DFL}}$. Table 13 presents those results for the two forecast selection techniques used. We see from that table that just over half the time the use of nine DFLs provides an RTA TE closer to zero.

TABLE 13

Aggregated Performance Comparison, on an Individual Flight Basis, Comparing the Number of Occasions That Nine DFLs Outperformed Four DFLs

Selection Technique	Nine Better Than Four (Counts)	Four Better Than Nine (Counts)	Mean of |TE| Improvements	Mean of |TE| Degradations
Optimized Headwind	74 (54%)	63 (46%)	1.8 s	1.1 s
Optimized Magnitude	73 (53%)	64 (47%)	1.4 s	1.1 s

One should then question how much did the use of nine DFLs degrade the performance for the other half of the flights? Were there occasions where the performance degraded significantly? It is found that, for the first question, the mean degradation is only 1.1 seconds regardless of which selection criteria was employed.

When the use of nine DFLs did improve the performance, it did so with an average of 1.4 seconds for the optimized wind magnitude approach and by 1.8 seconds for the headwind approach. The larger 1.8 second value could be explained by the fact that the standard deviation in RTA TE for this approach was larger than the other, and thus there is a greater change to create improvements.

Figure 34 shows the distribution of the improvements on a per flight basis. An improvement is a reduction of the |RTA TE|. This is formulated as

$$TE_{diff} = |RTA\ TE|_{9\ DFL} - |RTA\ TE|_{4\ DFL} \tag{4}$$

The more towards negative infinity the value, the greater the improvement. A positive value is a degradation. What would be desirable is to see that all the histogram bars are to the left of zero, meaning there is an improvement on each run. One of the things that can be seen from this figure is that even if there is degradation in performance, there does not appear to be any cases of significant degradation, though we do see some cases with as much as an additional 6 seconds of RTA TE.

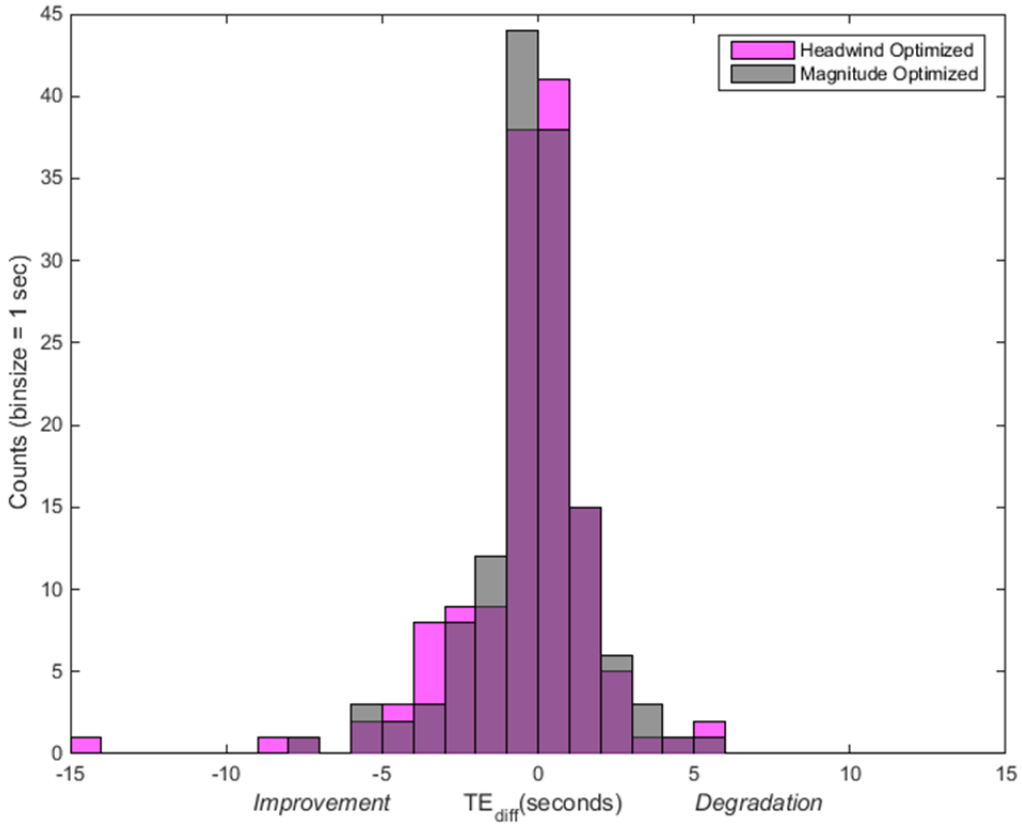

Figure 34. Histogram of RTA time error improvements using nine DFL instead of four DFL. It's desirable that all the counts are left of zero indicating that there is an improvement with each execution.

The second way to look at these data is driven by insight of what is the control strategy of this particular FMS. The FMS is programmed with an RTA TE tolerance with the allowable minimum being 6 seconds. This is also the value employed in these experiments. The FMS will implement speed changes if its estimated RTA TE is outside its active tolerance range. Internally, the actual controlled tolerance inside the FMS varies throughout the course of a flight, starting out much larger than the programmed tolerance, but by the time the aircraft approaches the RTA fix, it is trying to control to the specified tolerance. In fact, we know from the manufacturer that it is actually controlling to one half the set tolerance.

The utilization of a tolerance means that there is a deadband where additional control is not applied, and as far as the control system is concerned, any RTA TE within the deadband equates to no error. As such, to evaluate the question as to whether a particular forecast selection strategy or the choice of the number of descent forecast improve the *system's* performance, you must treat all RTA TEs within the actual controlled tolerance as zero error.

We reevaluate the individual flight performance as done earlier, but this time treating a change in RTA TE equal to or within the actual tolerance (±3 seconds) as zero error. We see in Table 14 that roughly 40% of the time there was no change in performance regardless of which forecast selection technique was used nor whether four or nine DFLs were used. Surprisingly, only 1/3 of the experiments showed improved time errors when nine DFLs were used, which is far less than the expected 90% from DL-1. Thus the remaining 1/3 performed better, if only marginally, if only four DFLs were used. A visual representation of this is given in Figure 35.

TABLE 14

Aggregated Performance Comparison Performed on an Individual Flight Basis, Comparing the Number of Occasions That Nine DFL Outperformed Four DFL Accounting for the Control Objective Deadband

Selection Technique	Nine Better Than Four (Counts)	No Change (Counts)	Four Better Than Nine (Counts)	Mean of |TE| Improvements	Mean of |TE| Degradations
Optimized Headwind	49 (36%)	54 (36%)	34 (28%)	2.3 s	1.1 s
Optimized Magnitude	39 (30%)	57 (42%)	41 (28%)	1.7 s	1.0 s

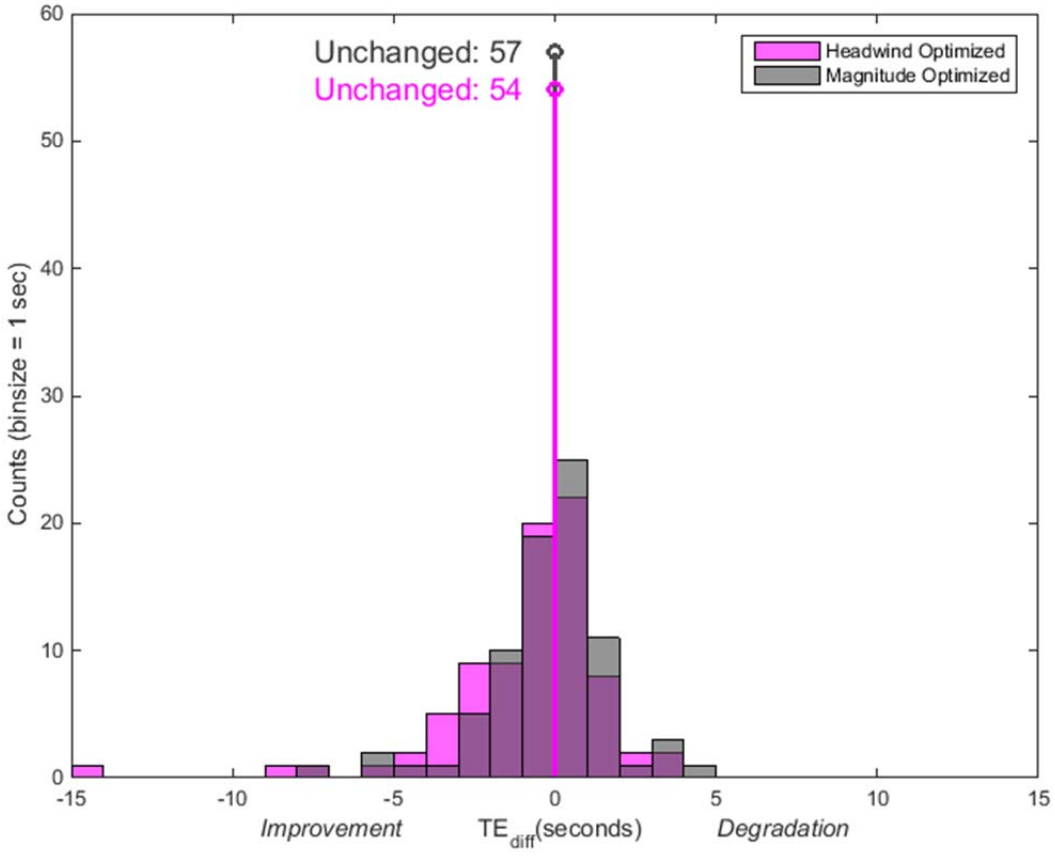

Figure 35. Histogram of RTA time error improvements using nine DFL over four DFL when accounting for control objective deadband. It is desirable that all the counts are left of zero, indicating that there is an improvement with each execution.

5.7 CONCLUSIONS

The systems development and research conducted in this phase of work was performed to evaluate questions germane to time-based NextGen applications including 4D-Trajectory Based Operations applications and required time of arrival operations.

The increased fidelity of the simulation components has begun to indicate the potential RTA performance future FMS systems could obtain given current weather prediction models. For the system tested, we find that there is a reduction in the standard deviation of the RTA time errors when an increased number (9 versus 4) of DFLs are used, though the improvement is quite small. For an individual flight, we find that using nine DFLs only provides reductions of RTA time errors in just over 50% of the simulated flights, and when it does, the improvements are on the order of 1.4 to 1.8 seconds. However, for

the remaining simulations, when there is no improvement by using more DFLs, the degradation is only 1.1 second on average.

The DFL selection algorithm is a factor on both the RTA time error bias and the magnitude of the standard deviation. The effect of different selection algorithms is more pronounced when using only four DFLs. The performance was nearly identical amongst the techniques when using nine DFLs, which would suggest that the wind profile could be sufficiently modeled with less DFLs.

With the provisioning of truth data as forecast information, we see a meaningful improvement in aggregated RTA performance for both the DFL selection algorithms studied. This is an indicator that improvements in forecast quality would be beneficial for future FMS systems, even if using an increased number of DFLs. However, there is no guarantee that the provision of more accurate forecast data will always provide better RTA performance on an individual flight due to model and control errors present in other components of the system.

It is difficult to make a sweeping conclusion on overall performance due to the limited number of samples available at the time of this writing. The current results indicate none of the RTA hypotheses could be supported, but this conclusion could change with the incorporation of additional simulations and at different airports.

6. SUMMARY AND RECOMMENDED NEXT STEPS

In this effort, a number of important improvements were made to the analysis infrastructure. These included doubling the number of FMS instances permitted to run in the WIAF from 20 to 40, reorganization of the simulation agent modules for improved infrastructure scalability and system management. The creation of MAFID, the Meteorological and Flight Information Database system, permitted the development and execution of high-fidelity reproductions of actuals flights. MAFID also allowed for the application and evaluation of the various forecasts that would have been available to the airline operating centers at the times of the flights.

A more rigorous assessment of wind forecast model accuracy was conducted by comparing HRRR model wind forecasts against MDCRS aircraft reports over a one-year period and four different U.S. airport regions. Results from trajectory-based wind forecast accuracy metrics such as headwind mean absolute error (MAE) and ETTF were stratified by a number of criteria, including forecast look-ahead time, location, altitude, flight phase, time of year, and time of day. A key finding from this study was relatively larger mean wind forecast errors and variability seen in the Northeast (EWR airport) region. It is hypothesized that this is in part due to the dynamic wind environment (e.g., frequent occurrence of vertical wind shear) of this region. Characterizing and determining the frequency of occurrence of wind environment scenarios (such as vertical wind shear) associated with large forecast errors, particularly in the Northeast region, and comparisons of GFS model forecasts against MDCRS, are important recommended follow-on activities. Additional suggestions for wind forecast model assessment are given in Section 3.7.

It has been shown that increasing the number of descent forecast levels in RTA operations can have a positive effect on overall performance, but the effect was smaller than expected. This is partially explained in that the overall performance of the systems modelled, which can be considered a relatively high-performance system already, is such that there is a limited area for improvement. An important question that remains is what is the effect of increasing the number of DFLs for other airframes and FMS systems? We recommend evaluating additional airframes and avionics suites to determine if additional and/or more accurate forecast information can improve the RTA performance of those systems. In particular, we recommend evaluating the effect on systems that incorporate two other popular FMS systems on two very popular airframes, the GE/Smiths FMS on the B737 and the Thales FMS on the A320.

Our results show that there is an effect on RTA performance based on the DFL selection technique, but this area has not been fully explored. Noting that there is an effect on the system evaluated raises the question if there would be a more pronounced effect on other systems. Consider the case of the GE/Smith FMS in a B737, which currently can only use three DFLs. The limited number of DFLs would likely make the system's RTA performance more susceptible to error as a function of reduced fidelity of the descent wind profile. To some degree, this system can be emulated within the existing infrastructure by

limiting the DFL selection to three levels. One could then explore the trade spaces on the DFL selection techniques (which specify the locations, altitudes, and times) and the available sources (age, forecast, or locally enhanced forecast). The existing infrastructure could also be modified to emulate the autothrottle behavior of the B737/FMS to add further realism to the emulation.

Whereas this work evaluated simulated descents to around 10,000 ft MSL, we recommend conducting an analysis on RTA performance to significantly lower altitudes. Though the initial incarnation of RTA operations were not developed to conduct operations to altitudes much below 10,000 ft, the results of this study suggest that the system evaluated could perform to lower altitudes. We believe simulated flights descending to at least the altitudes associated with most initial approach fixes and perhaps even to final approach fixes should be conducted. If it is shown that RTA systems can remain performant to these altitudes, these findings could have a significant effect on the concept and application of RTA in future 4D-TBO operations.

The vast majority of the STARs flown in these simulations had speed constraints on waypoints along the routes before the RTA fix location. In some cases, the first of several speed constraints started as high as FL330 (e.g., EAGUL6 at KPHX). Current RTA requirements require that all speed constraints along a route are to be respected. As such, there is limited speed control authority available to the RTA system to correct for errors. It is recommended that the effect of speed constraints on RTA performance be evaluated to inform future CONOPS and procedure design.

It has been shown that perfect forecast data, i.e., truth data, can improve RTA performance. To improve the accuracy of forecasts, it is suggested that data derived from aircraft operating locally (spatially and temporally) could be obtained as a potential source of information that could be used to augment existing forecasts. One potential means to obtain such information could be from the interrogation and interpretation of data available from aircraft operating in the vicinity who are outfitted with Mode-S EHS transponders. Such transponders could provide aircraft state data that can be used to estimate current wind conditions. It is recommended that this avenue be investigated to first determine the method (e.g., interrogation schedule) and effect (e.g., increased bandwidth usage) of collecting such data. Secondly, the use of such data must be evaluated in terms of how it could be leveraged to improve or augment forecasts, as well as how it could be employed for operations other than RTA, such as Interval Management or real-time wake vortex mitigation validation.

It is recommended that the evaluation of RTA performance be expanded to a larger set of airports and time frames to include the effects of arrival route design and local meteorological phenomena. An example airport of interest is KEWR, which is in a region where the HRRR model appears to have overall reduced forecast accuracy. Similarly, it is recommended that the GFS (Global Forecast System) model also be evaluated, as this is currently the mostly widely used forecast by commercial carriers for flight planning.

APPENDIX

HRRR DATA PROCESSING AND STORAGE FOR MAFID

Aircraft operations of 4D-TBO rely on weather forecast data to assist in achieving RTA objectives. The current phase of analysis for RTA simulation utilizes High-Resolution Rapid Refresh (HRRR) data as its forecast source. Continuous ingesting of this data in its original format raises a challenge as each HRRR data file in its provided GRIB2 format is on average 350 MB in size. A single day of HRRR data includes around 376 individual files and utilizes approximately 130 GB of file system space. The nature of the analysis in this work required fast access to forecast data. The limitations of storage and the time taken to load any significant amount of HRRR forecast data in its original format was not conducive to this requirement.

HRRR forecast files superimpose grids of weather forecast data covering CONUS, as seen in the example of the HRRR product "V-component of wind" in Figure A-1.

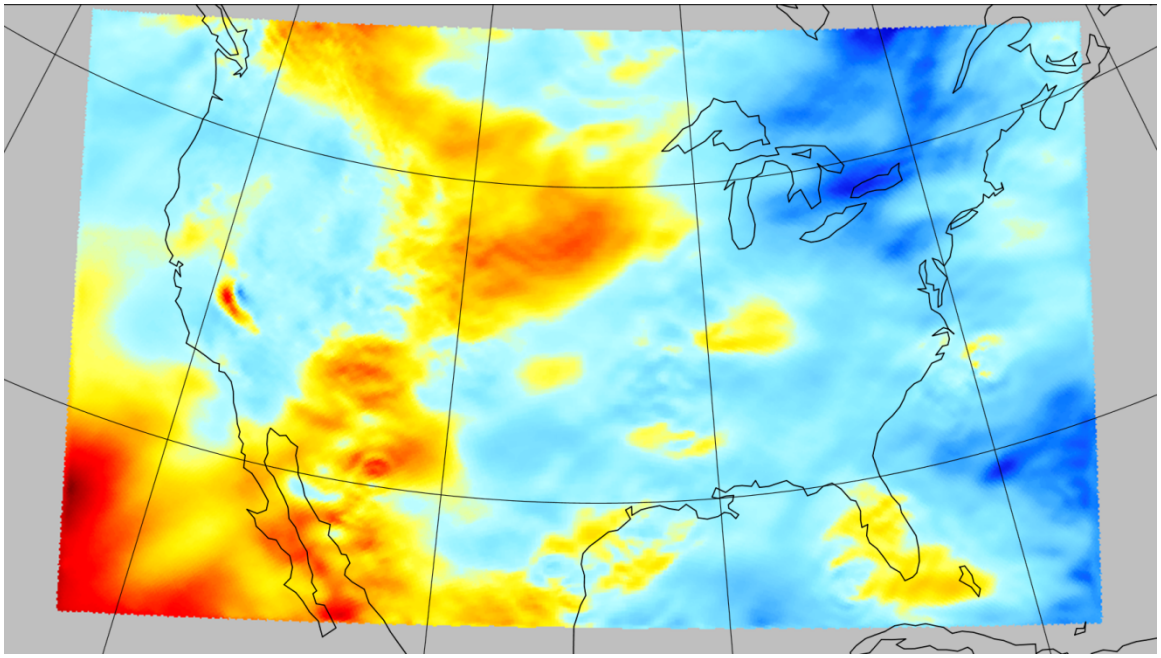

Figure A-1. Lambert conformal view of V-component of wind with heat map from HRRR sample.

The original HRRR data files contain many products (distinct data types) that are not needed for the current program objectives. For example, “Snow depth @ Ground or water surface” or “Composite reflectivity @ Entire atmosphere” are samples of the many produced data products that we do not need to retain in our data store. Data products related to wind, temperature, and pressure are the primary elements we must retain along with a set of additional product we are collecting for potential future analysis. This “pruning” of the HRRR forecast products significantly reduce the size of the resulting files. After removing the unneeded products, the average HRRR data file is reduced from 350 MB to approximately 77 MB.

As seen in Figure A-1, the HRRR data files represent forecast data for the entire CONUS in lambert conformal projection. As part of the data processing stream, these data are reprojected and stored in a Plate Carre (lat/Lon) projection, which is a more expeditious base form to use for the web-service forecasting service that is geared to aviation. Also, the full CONUS data set is broken into smaller tiles. The reduced tiles size makes for quicker loading and reduced caching for the forecasting service. Figure A-2 shows the tiling of the subdivide HRRR data as stored in the database to ensure there is no data loss; the original data is slightly oversampled in the reprojected data. The additional sampling and the effect of tiling do, however, have a negative effect on the data compression, and the net total storage for an individual product is greater than required when in its original form.

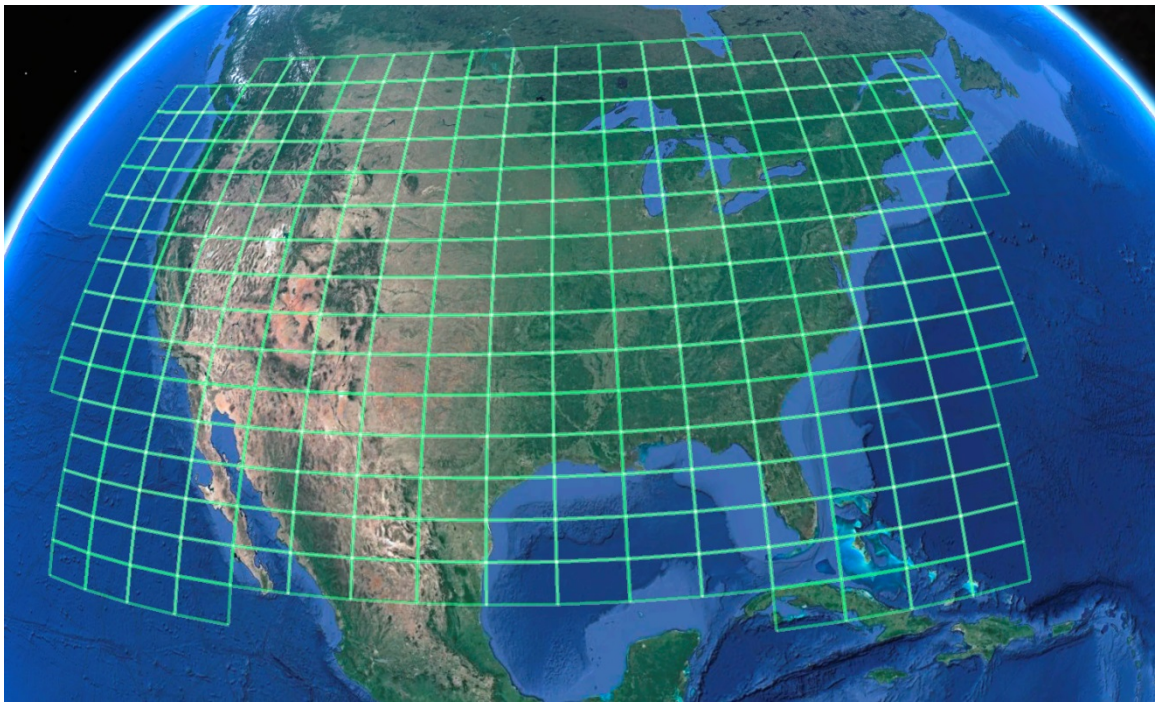

Figure A-2. Generated data tiles after latitude/longitude reprojection and subdividing.

Each hourly HRRR forecast collection has valid forecast out to 18 hours from the publication time. The range of forecast information is not required for our analysis, so we only retain forecast information out to eight hours. This, along with the product pruning, reduces the total daily storage from 130 GB to 24 GB. As noted above, the original HRRR data set for a single day was in excess of 130 GB in size. After processing and compression of the tiled HRRR files, the entire day of data requires approximately 24 GB of file system space.

WEATHER FORECAST DATA SERVICE

To enable a number of wind forecast related activities, a forecast weather service has been designed based on a database and file-system back-end. Currently implemented for HRRR, the architecture is extensible to include other types of data and other data containers (currently only GRIB containers are supported). Using HRRR data, the queryable forecast weather service is capable of returning the interpolated forecast wind speed, direction, and air temperature at a particular (lat, lon, pressure altitude, time), as well as surface pressure and surface winds. For higher access performance, the CONUS-wide HRRR tiles were retiled to smaller size. These are found during the query process, either contained in memory of the local cache, or on disk, pointed to via a database query. To forecast the value of a single point in (lat, lon, alt, time), it is necessary to use two tiles that bound the spatial point within time by belonging to forecasts describing nearest times before and after the requested time. The information from these tiles is then interpolated to produce the final data product.

The 4D-TBO Winds program presents a number of challenges to supporting infrastructure that dictate its architecture and necessitate particular approaches to be taken in its design. In particular, the requirements were

1. Simplify access to sophisticated weather data products for multiple consumers
2. Minimize development efforts required to create weather product client applications
3. Provide performant client access to forecast weather data products
4. Allow flexibility in use of different types of weather data products, locations, and storage formats in the near future (as HRRR is transitioning to a different storage mechanism)

The design of the weather (WX) server was defined by these requirements. The WX server, a dedicated web service, provides a simplified front end to a significant amount of activity that must be carried out in order to obtain a final data product. A request for weather information given particular space/time coordinates entails locating the data and performing interpolation across space and time. Raw data access is onerous, error handling is de rigueur, and mathematics must be performed on 16 nearby data products to obtain a single interpolated value. Thus, it makes sense to perform this error-prone activity as few times as possible, in a centralized fashion.

A reduction in development time was experienced due to the use of standardized access methods (HTTP REST) and data marshalling formats (JSON) that helped decouple application-specific data processing from generic data request and handling activity. The access methods and formats were also chosen to be usable across multiple platforms and programming languages, so interoperability was more easily achieved, and development could take place easier in languages that were more suited to the task – a follow-on effect of proper architecture.

Leveraging centralized in-memory storage of weather data and thus reducing the overhead costs of queries and loads across multiple distributed simulation runs provides a significant performance gain. Trajectory-based weather data are particularly well-suited for this type of data infrastructure, as they follow predictable data access patterns.

In addition to the currently supported file formats and containers, other types of containers (e.g., CDF+/HDF) and forecasts (GFS) are under consideration and could be added with minimal development effort.

Figure A-3 provides a top-level view of the WX data server architecture. A user request enters the system at the user interface and is validated against the tiles that exist in the local memory of the system by the cache processor. If the tiles representing the spatial neighborhood and bounding the nearest temporal extents of the requested spatiotemporal point are currently stored in the cache, they are simply passed to the space/time interpolation module, and the interpolated product is then returned to the user. This is the most efficient use case of the system. It can occur frequently, since, even though trajectories typically span multiple geographic locales and temporal slices, continuous slices of a trajectory are frequently composed of several data points contained in a single tile, with multiple tiles connected to form a complete trajectory slice. Once a single pair of tiles is loaded, points contained inside that tile do not cause a reload penalty with the current caching scheme. Future work in the caching scheme will introduce garbage collection methods that rely on a least-recently-used or a least-frequently-used scheme to discard cached data.

If the tile is not found in the cache, a database query is constructed to attempt to locate a tile with the necessary information. The query returns a file location and name, which are then loaded from a network file system by the data object loader. Once loaded, the tile is then stored in the cache, and also forwarded to the space/time interpolation module for computation.

Space/Time Interpolation performs bilinear interpolation in lateral space, then vertically, and then in time. This is done on a per-data-product basis, with a notable exception of the wind fields, in which the northing and easting wind components are interpolated separately and then combined to form wind magnitude and direction.

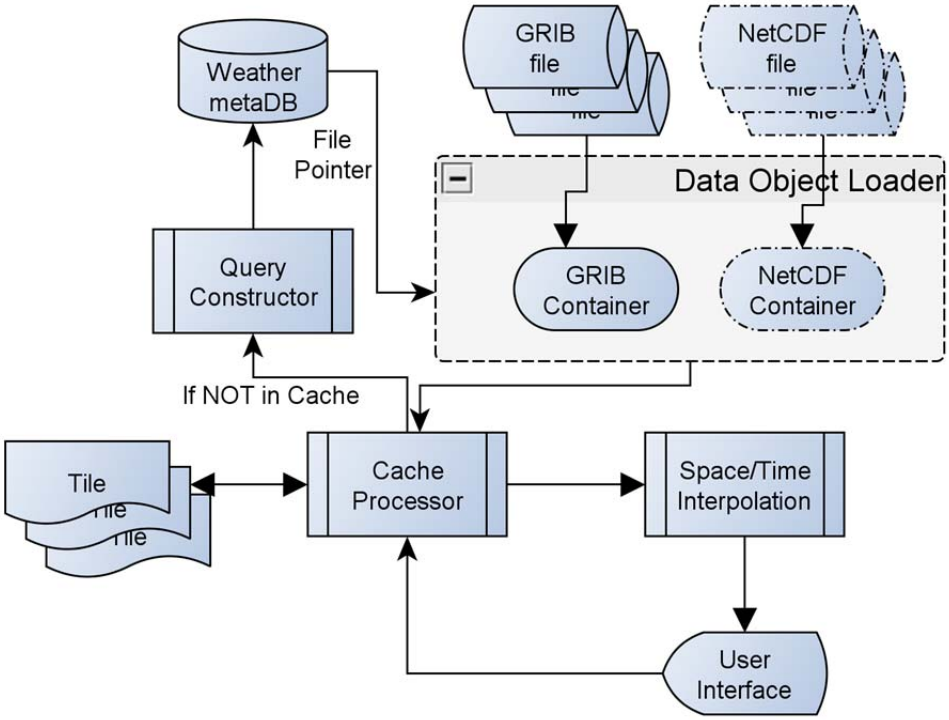

Figure A-3. WX data server architecture.

This page intentionally left blank.

GLOSSARY

4D-TBO	4D-Trajectory Based Operations
A4A	Airlines for America
ABP	Achieve by Point for interval management procedures, equivalent to meter fix in TOAC procedures
ACARS	Aircraft Communications Addressing and Reporting System
AGL	Above Ground Level
AHRS	Attitude Heading Reference System
AIS	Aeronautical Information Services
AOC	Airline Operations Center
ARAM	Aircraft Reported Atmospheric Model
ASDE-X	Airport Surface Detection Equipment, Model X
ASG	Assigned Spacing Goal
ATC	Air Traffic Control
ATM	Air Traffic Management
ATTF	Actual Time to Fly
CIFP	Coded Instrument Flight Procedures
CONOPS	Concept of Operations
CONUS	Contiguous United States
CTA	Controlled Time of Arrival
DFL	Descent Forecast Level
EHS	Enhanced Surveillance
ERAM	En Route Automation Modernization
ESRL	NOAA Earth System Research Laboratory
ETTF	Estimated Time to Fly
EWR	New York Newark Liberty International Airport
FAF	Final Approach Fix
FIM	Flight-deck Interval Management
FL	Flight Level, an altitude level of constant atmospheric pressure relative to international standard sea level pressure of 29.92 inches of mercury. Every flight level is stated in hundreds of feet, with the last two zeros removed.
FMS	Flight Management System

FY	Fiscal Year
GE	General Electric
GFS	Global Forecast System Model, U.S. NOAA National Centers for Environmental Prediction (NOAA/NCEP XE "NCEP" \t " NOAA National Centers for Environmental Prediction ") wind forecast model with 25 km spatial resolution used as basis for many airline flight planning products
GIM	Ground-based Interval Management
GSD	NOAA Global Systems Division
HRRR	High-Resolution Rapid Refresh Model, U.S. NOAA/NCEP/ESRL wind forecast model with 3 km spatial resolution used as basis for FAA high-resolution weather forecasting products
HW	Headwind
IAP	Instrument Approach Procedure
IAS	Indicated Air Speed
IM	Interval Management
KATL	Hartsfield-Jackson Atlanta International Airport
KBOS	Boston Logan International Airport
KCLT	Charlotte/Douglas International Airport
KDEN	Denver International Airport
KDFW	Dallas/Fort Worth International Airport
KIAH	George Bush Intercontinental Airport
KJFK	John F. Kennedy International Airport
KMDW	Chicago Midway Airport
KMEM	Memphis International Airport
KSDF	Louisville International Airport
LL	Lincoln Laboratory
MADIS	Meteorological Assimilation Data Ingest System
MAE	Mean Absolute Error
MAFID	Meteorological And Flight Information Database
Magn	Magnitude
MCDU	Multifunction Control Display Unit
MDCRS	Meteorological Data Collection and Reporting System, Meteorological Data Collection and Reporting System
Meter fix	Location where aircraft is targeting to get to by the CTA/RTA is controlled to by FMS in TOAC procedures

MSL	Mean Sea Level
NAS	National Airspace System
NATCA	National Air Traffic Controllers Association
NavDB	Navigational Database
NCEP	NOAA National Centers for Environmental Prediction
NM	Nautical Mile (1853 meters or 6080 ft)
NOAA	National Oceanic and Atmospheric Administration
ORD	Chicago O'Hare International Airport
PHX	Phoenix Sky Harbor International Airport
PPH	Pounds per Hour
RMS	Root Mean Square
RMSVD	Root Mean Square Vector Difference
RMSVE	Root Mean Square Vector Error
RTA	Required Time of Arrival function of an FMS which manages aircraft speed in an attempt to comply with CTA at the meter fix
RTA TE	Required Time of Arrival Time Error (actual time of arrival at meter fix relative to the target time)
RTCA	Radio Technical Commission for Aeronautics
RUC	Rapid Update Cycle, U.S. NOAA/NCEP wind forecast model, predecessor of RAP
SC-186	RTCA Special Committee for Automatic Dependent Surveillance – Broadcast
SC-206	RTCA Special Committee for Aeronautical Information Services Data Link
SC-227	RTCA Special Committee for Standards of Navigation Performance (including TOAC)
SC-214	RTCA Standards for Air Traffic Data Communications Services
SFO	San Francisco International Airport
SG-7	Sub group of RTCA Special Committee for Aeronautical Information Services Data Link, Wind Information
SID	Standard Instrument Departure
STAR	Standard Terminal Arrival Route
TFMS	Traffic Flow Management System
TMC	Thrust Management Computer
TOAC	Time of Arrival Control
TOD	Top of Descent (end of cruise, start of descent)

TRACON	Terminal Radar Approach Control
TTF	Traffic to Follow
VNAV	Vertical Navigation
WIAF	Wind Information Analysis Framework
WTIC	Weather Technology in the Cockpit
WX	Weather

REFERENCES

1. SC-227, "Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation DO-236C," RTCA, Washington DC, 2014.

2. Reynolds, T.G., Y. Glina, S.W. Troxel, and M.D. McPartland, "Wind Information Requirements for NextGen Applications Phase 1: 4D-Trajectory Based Operations (4D-TBO)," Project Report ATC-399, MIT Lincoln Laboratory, 2013.

3. Glina, Y., T.G. Reynolds, S. Troxel, and 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.

4. Reynolds, T.G., S. Troxel, Y. Glina, and M. McPartland, "Establishing Wind Information Needs for Four Dimensional Trajectory-Based Operations," *1st International Conference on Interdisciplinary Science for Innovative Air Traffic Management*, Daytona Beach, FL, 2012.

5. Reynolds, T.G., P. Lamey, M. McPartland, M. Sandberg, T. Teller, S. Troxel, and Y. Glina, "Wind Information Requirements for NextGen Applications Phase 2 Final Report: Framework Refinement & Application to 4D-Trajectory Based Operations (4D-TBO) & Interval Management (IM)," Project Report ATC-418, MIT Lincoln Laboratory, March 2014.

6. Edwards, C., McPartland, M.D., Reynolds, T.G., Sandberg, M., Teller, T., and Troxel, S., "Wind Information Requirements for NextGen Applications Phase 3 Final Report," Project Report ATC-422, MIT Lincoln Laboratory, October 2014.

7. NOAA MADIS ACARS website, http://madis.noaa.gov/madis_acars.shtml

8. Schwartz, B.E., S.G. Benjamin, S. Green, and M. Jardin, "Accuracy of RUC-1 and RUC-2 Wind and Aircraft Trajectory Forecasts as Determined from ACARS Observations," *Weather and Forecasting*, 15, June 2000.

9. Robert, E. and D. De Smedt, "Comparison of Operational wind Forecasts with Recorded Flight Data," *Tenth USA/Europe Air Traffic Management Research and Development Seminar* (ATM2013), 2013.

10. Cole, R.E. and Kim. S.K., "A Study of Time-to-Fly Estimates for RUC and ITWS Winds," *Ninth Conference on Aviation, Range, and Aerospace Meteorology* (ARAM), Orlando, FL, Amer. Meteor. Soc., 2000.

11. Ahmad, N., Barmore, B., and Swieringa, K., "Wind Information Uplink to Aircraft Performing Interval Management Operations," *15th AIAA Aviation Technology, Integration, and Operations Conference*, AIAA Aviation (AIAA 2015–3398), 2015.

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