note technical note tech

Evaluation of an ERAM Prototype to Improve Restriction Modeling by Refining Altitude Transition Rate Logic

Brian S. Schnitzer, General Dynamics Information Technology

Andrew Fabian, FAA ANG-C41

Christina M. Young, PhD, FAA ANG-C41

Chu Yao, FAA ANG-C41

September 2013

DOT.FAA/TC-TN13/42

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

U.S. Department of Transportation
Federal Aviation Administration

William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT.FAA/TC-TN13/42		

4. Title and Subtitle	5. Report Date
Evaluation of an ERAM Prototype to Improve Restriction Modeling by Refining Altitude Transition Rate	September 2013
	6. Performing Organization Code ANG-C41
7. Author(s) Brian S. Schnitzer, Andrew Fabian, Christina M. Young and Chu Yao	8. Performing Organization Report No. DOT.FAA/TC-TN13/42
9. Performing Organization Name and Address U. S. Department of Transportation Federal Aviation Administration, William J. Hughes Technical Center Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address U. S. Department of Transportation Free Flight Phase 1 Program Office, AOZ-200 Conflict Probe/URET Product Team Washington, D. C. 20590	13. Type of Report and Period Covered Technical Note 14. Sponsoring Agency Code DOT

15. Supplementary Notes

The authors identified above represent the following organizations: Brian S. Schnitzer with GDIT; Andrew Fabian, Christina M. Young and Chu Yao with FAA ANG-C41

16. Abstract

The Federal Aviation Administration (FAA) is currently implementing a number of improvements to the National Airspace System (NAS) in the United States under a multi-agency initiative called the Next Generation Air Transportation System (NextGen) Program. The Separation Management and Modern Procedures Project is a NextGen initiative and its objective is to implement the En Route Automation Modernization (ERAM) strategic conflict probe on the radar controller display utilizing ERAM's Trajectory Modeling (TM) and Conflict Probe (CP) sub-systems. The FAA Air Traffic Organization's En Route Program Office (ATO-E) has employed the FAA's Concept Analysis Branch (ANG-C41) to conduct a series of independent evaluations on prototype enhancements to the TM and CP sub-systems and has contracted the prime contractor of ERAM, Lockheed Martin, under FAA Task Orders 45 and 51 to develop these prototypes within the ERAM architecture. This paper details an experiment that consists of simulated runs using the ERAM system with and without a prototype enhancement designed to improve restriction modeling and climb/descent modeling. Recorded data from real flights in Chicago (ZAU) and Washington (ZDC) centers are used to generate experimental scenarios which provide realistic air traffic scenarios, for a total of 4 scenario runs. For each scenario, prototype Trajectory Modeler (TM) and Conflict Probe (CP) performance is compared to that of the baseline scenarios. In both ZAU and ZDC scenarios, slight improvements are indicated with regard to the CP. There is no indication of trajectory improvement or degradation. Since the net effects are positive, particularly for false alerts, the prototype is considered to be an improvement to the system.

17. Key Words	18. Distribution Statement
Separation Management, Conflict Probe, En Route Automation Modernization, ERAM, Trajectory Modeling, Restriction	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	27	

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

[THIS PAGE IS INTENTIONALLY LEFT BLANK]

Executive Summary

The Federal Aviation Administration (FAA) is currently implementing a number of improvements to the National Airspace System (NAS) in the United States under a multi-agency initiative called the Next Generation Air Transportation System (NextGen) Program. The NextGen operational concept envisions a future air traffic environment managed by aircraft trajectory with advances in ground automation like the conflict probe. The Separation Management and Modern Procedures Project is one of these NextGen initiatives and its objective is to implement the En Route Automation Modernization (ERAM) strategic conflict probe on the radar controller display. The strategic conflict probe utilizes ERAM's Trajectory Modeling (TM) and Conflict Probe (CP) sub-systems to notify air traffic controllers when aircraft will violate separation standards as much as 20 minutes in the future. The FAA's Air Traffic Organization's En Route Program Office (ATO-E) contracted the prime contractor of ERAM, Lockheed Martin, under FAA Task Orders 45 and 51 to develop these prototypes within the ERAM architecture so the FAA may evaluate their efficacy. ATO-E has employed the FAA's Concept Analysis Branch (ANG-C41) to conduct a series of independent evaluations on performance enhancements to the TM and CP sub-systems.

This paper describes an experiment designed to study a prototype enhancement that alters how the track-weighed altitude transition rate factor (TWF) is calculated. A component of the TM, this algorithm is utilized when intent information indicates that a flight will climb or descend in the future. It was determined that the legacy method of calculating the TWF, which applies constraints on the altitude transition rate, occasionally causes the TM to build a trajectory that does not meet a restriction even though subsequent track data indicates that the aircraft does indeed meet the restriction. Modifications to the TWF were introduced in an effort to correct this issue with the goal of improving trajectory accuracy. In the prototype algorithm, the TWF is not constrained based on previous descents/climbs. The TWF is reset to unity when a level trajectory segment is predicted and then calculated again for the next altitude transition segment, removing any bias from previous descents or climbs.

The experiment performed by CA consists of simulated runs using the ERAM system. The TM and CP performance of these treatment runs are compared to that of their respective baseline runs, which represent the current state of the live ERAM system. Two baseline runs are used in this experiment; one from the Chicago (ZAU) Air Route Traffic Control Center (ARTCC) and one from the Washington (ZDC) ARTCC. These two runs were originally from 2010 recordings of live traffic during peak hours. The traffic data is then time-shifted to induce conflicts, which are useful for testing the CP because they rarely occur in recorded data.

In this experiment the TWF prototype resulted in a slight improvement in False Alert (FA) counts without causing any degradation to Late Alerts (LA), Missed Alerts (MA), or Warning Time (WT). Trajectory accuracy did not show any significant changes from a practical standpoint. However, LM indicated that there are restrictions associated with incorrect STARS, restrictions with improper qualifiers ("AT_OR_BELOW" instead of "AT_ALTITUDE"), and some restrictions may no longer be used at all. Therefore, adaptation data as it pertains to restrictions needs to be validated in order to better gauge the scenario-wide benefit of the TWF prototype. Since the TM utilizes restriction information when predictions are made about where an aircraft will fly, and given the scope of the experiment, it is not known how many flights suffered degradation in trajectory accuracy due to inaccurate restriction data in the ARTCC adaptation. Since the net effects were positive, particularly for false alerts, the upgrade is considered to be an improvement to the system and is recommended for deployment from an engineering standpoint.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background to Study

This activity supports the development of mechanisms and methods to improve the Trajectory Modeling (TM) and Conflict Probe (CP) performance within ERAM. It consists of a joint effort between Lockheed-Martin (LM) and the Concept and Analysis (CA) Team to analyze the effectiveness of a new prototype enhancement. The CA Team received baseline and treatment scenarios from LM in an effort to compare the effect that the prototype has on the TM and CP metrics.

1.2 Prototype Enhancement

[Bentz and McKay, December 2012] found that the TM occasionally predicts that a restriction cannot be met, even in cases where subsequent track data indicates that the aircraft does indeed meet the restriction. This prediction results in the TM building inaccurate trajectories, which necessarily lead to some inaccurate CP predictions. Results suggested that a change in how trajectories are modeled should be implemented. A new track weighted climb/descent factor (TWF) was introduced by LM as an effort to solve this problem. This prototype enhancement would require code enhancements to the current version of ERAM, and would allow for a resetting of the TWF following level trajectory segments, resulting in a less constrained predicted climb/descent rate for the TM. This study analyzes the effects of this prototype enhancement on both the Trajectory Modeling (TM) and the Conflict Probe (CP).

1.3 Scope of Study

This document reports on the results of an experiment limited to two nine-hour traffic samples. One scenario was collected on February 11, 2010 from the Chicago Air Route Traffic Control Center (ZAU). The other was collected on April 30, 2010 from the Washington Air Route Traffic Control Center (ZDC). To induce conflicts between aircraft and for evaluation purposes only, the data sample was time-shifted using the methodology documented in [Paglione, 2003].

All of the analyses in this document were performed on a time-shifted scenario. Currently, the metrics available for analyzing performance require a time-shifted scenario to be used in order to generate actual loss of separation that would not occur under normal circumstances. This time-shifting can create some events that the conflict probe will never encounter in a live system. As a result, the reader should be careful not to take any numbers presented in this document out of context, as the simulated scenarios should only be compared to other scenarios in this study. All numbers presented in this document should be used only for comparison to other numbers included in this document, unless otherwise noted. The False Alert, Late Alert, and Missed Alert rates, as well as the warning time values presented in this document do not reflect the actual values of the live ERAM system and should not be considered as such. Because of this, most of the values presented in this document are in the form of percentage change from the baseline results. Though some raw numbers may be presented, they should be considered only in the context of this document.

1.4 Document Organization

This technical note is organized in the following sections: Section 1 provides background for the study and a brief description of the prototype enhancement being analyzed in this study. Section 2 defines the experiment performed. Section 3 describes the analyses that were performed to evaluate the Trajectory Modeling (TM) performance and the Conflict Probe (CP) performance. Finally, Section 3.4.4 wraps up the conclusions of the performance analyses and makes recommendations based on the findings.

2 Description of Experiment

One of the most powerful inferential statistical approaches is the design, implementation, and synthesis of experiments. Experiments are performed by most researchers and scientists in practically all disciplines. An input stimulus is entered into a process with a set of controllable factors. The uncontrollable factors are not easily manipulated, but through experimental design techniques such as blocking and randomization can be removed from the experiment. The output response variables are the dependent variables of the experiment. They are often determined by application of a metric or measured by a sensor device.

Table 1. Processing Steps for the Experimental Analysis

Step	Description
1 – Problem Definition	Define the problem statement
2 – Design of Experiment	Design the experiment – The factors, levels of the factors, response variables to be run, and the model to be used for analysis are defined.
3 – Execute Experiment	Execute the experiment and prepare output data – The system is configured for the experimental runs defined by the design, runs executed, and resulting output data is processed for input into model
4 – Implement Model	Implement statistical model defined by the experiment.
5 – Model Results	Examine the results of the model and discuss factor effects
6 – Synthesize Impact	Synthesize overall results from the model and publish conclusions.

There are many purposes for performing an experiment. For this study, the objective of designing and executing an experiment is to establish (1) whether the factor in question shows a statistically significant effect on the ERAM system's performance, and (2) the relative sizes of the determined significant effects. From designing the experiment to concluding on its results, a series of processing steps should be performed as identified in Table 1. The first two steps are described in this section, which documents the plan for the experimental analysis. The last four steps are described in Section 3 and Section 3.4.4, which present the results by documenting the actual execution and analysis of the experiment.

Since there is only one factor manipulated in this study, this experiment is much simpler than previous experiments performed by Concept Analysis (CA). The purpose of this experiment is to evaluate whether the new track weighted descent factor provides a statistically significant improvement to the performance of the Trajectory Modeler (TM) and/or the Conflict Probe (CP) as compared to the legacy track weighted descent factor when considering restriction modeling. Legacy-prototype is a binary factor, and thus only one baseline and one experimental run is needed per scenario.

2.1 Definition of the Problem Statement

It must be determined if this prototype can provide a significant improvement to TM or CP. TM performance is measured in terms of Trajectory Error metrics, while CP performance is measured in False Alert, Late Alert, and warning time performance, all of which can vary separately. Low Trajectory Error metrics and Low False Alerts, low Late Alerts, and high warning time are the desired qualities of CP performance. The prototype covered in this study is not specifically intended to improve CP performance, but to address observed issues with how trajectories are built when restrictions are on the flight plan. For this study, the problem statement is expressed as follows:

> *Through a set of purposeful runs of ERAM, input with time-shifted test traffic scenarios from ZAU and ZDC, the experiment shall determine the effect, if any, that an enhanced algorithm for determining track weighted descent factors has on both scenario-wide trajectory and conflict prediction accuracy performance as well as on individual flights.*

A significant change, whether it is improvement or degradation, is defined as a change in the respective metric (False Alerts, Late Alerts, or warning time) that is greater than the confidence intervals of the statistical model. These confidence intervals are discussed in the next section.

2.2 Design of Experiment

In all analyses performed in this study the baseline (BL) used contains all of the settings of the current live system, plus the trajectory lateral modeling enhancements defined under Function Area 32 ap1 [McKay, 2011]. The result of the analysis on the FA32 ap1 was a recommendation for addition of those enhancements to the trajectory modeler. BL1 is the baseline run for ZAU (TWF is off) and BL2 is the baseline for ZDC (TWF off) . Since this experiment manipulates only one binomial factor, a full factorial design can be performed. Prototype enhancements are binary factors, either running or not running, and result in only two experimental runs, one for each scenario. These Track Weighted Factor (TWF) runs are also shown in Table 2.

Table 2. Runs for the experiment

Run	ARTCC	TWF
BL1	ZAU	Off
TWF1	ZAU	On
BL2	ZDC	Off
TWF2	ZDC	On

Both the baseline and treatment use the same parameters for trajectory modeling, conformance bounds, conflict detection, and likelihood settings as the currently fielded version of ERAM. These parameter settings are described in Table 3.

Table 3. Settings of the baseline run (BL) used for this experiment.

Lat TM	Lat CD	Lon TM	Lon CD	Likelihood
2.5	2.5	1.5	1.5	10|20

3 Performance Evaluation

The performance evaluation analyses used in this study are similar to those used in previous Concept Analysis experiments [Crowell et al, December 2011b]. The metrics used are those described in the documentation of Integrated Experiment 1 [Crowell et al, December 2011a]. For this experiment, trajectory accuracy is considered only in the vertical dimension and Conflict Probe analysis is virtually identical to previous experiments. An analysis novel to CA, a brief overview of restriction accuracy, is performed as well.

3.1 Trajectory Analysis

Analysis was performed on the experimental data to determine the effects that the TWF enhancement has on trajectory accuracy. The null hypothesis is as follows:

> *A statistically significant improvement in trajectory accuracy is not observed with addition of the prototype enhancement T*rack-Weighed Altitude Transition Rate Factor

A significant performance improvement is defined as a statistically significant reduction in any of the trajectory error metrics used by Concept Analysis to measure trajectory accuracy. Since the TWF enhancement by design only affects the vertical component of the TM, only the Vertical trajectory error will be evaluated here. The consideration of whether any improvement in trajectory accuracy is of practical or operational significance is also important. A conservative threshold of a 100 ft minimum change in Vertical trajectory accuracy will be used in this paper.

Table 4 shows the average absolute difference in vertical trajectory error for all flights in the ZAU scenario. Negative mean differences represent a reduction in error when the TWF enhancement is included. Table 5 shows the average absolute difference in vertical trajectory error only for flights in the ZAU scenario that had active restrictions of any sort on their flight plan trajectory.

Table 4 . ZAU scenario: overall difference in vertical trajectory error.

Look Ahead (sec)	Mean Difference (ft)	SD	N	Significance[1]
0	-0.5	23	2234	NS
300	5.5	207	2201	NS
600	9.5	218	2024	NS
900	7.3	129	1561	p < .05
1200	6.0	68	1051	p < .01

[1]For significance, "NS" means not significant and the null hypothesis cannot be rejected. Both $p<0.05$ and $p<0.01$ indicated significance and the null hypothesis can be rejected. Rejection of the null hypothesis suggests that the TM with TWF is different from the baseline.

Table 5. ZAU scenario: difference in vertical trajectory error for flights with restrictions.

Look Ahead (sec)	Mean Difference (ft)	SD	N	Significance
0	-1.8	28	1072	p < .05
300	11.4	292	1060	NS
600	17.3	304	984	NS
900	13.0	174	752	p < .05
1200	13.5	94	423	P < .01

Table 6 shows the average absolute difference in vertical trajectory error for all flights in the ZDC scenario. Table 7 shows the average absolute difference in vertical trajectory error only for flights in the ZDC scenario that had active restrictions of any sort on their flight plan trajectory.

Table 6. ZDC scenario: overall difference in vertical trajectory error.

Look Ahead (sec)	Mean Difference (ft)	SD	N	Significance
0	-0.5	19	2664	NS
300	-5.3	126	2543	p < .05
600	0.1	161	2301	NS
900	3.0	246	2052	NS
1200	0.2	231	1742	NS

Table 7. ZDC scenario: difference in vertical trajectory error for flights with restrictions.

Look Ahead (sec)	Mean Difference (ft)	SD	N	Significance
0	-0.8	20	2104	NS
300	-7.4	128	2020	p < .05
600	-2.5	111	1816	NS
900	-2.2	118	1608	NS
1200	-5.9	100	1352	p < .05

The average absolute vertical trajectory error was calculated for each flight in both the baseline and the treatment scenarios. A matched pairs test on the difference in vertical trajectory error between the same flights in the baseline and treatment scenarios was performed. Both the ZAU and ZDC scenarios show some statistical difference in vertical error metrics between their respective baseline and treatment scenarios. This shows up as very slight improvement (negative means) or degradation (positive means) in accuracy, and is virtually meaningless from a practical/operational standpoint, as the mean difference never goes beyond 20 ft. This does not preclude individual cases where an improvement in vertical error occurs, however.

3.2 Conflict Probe Analysis

Analysis was performed on the experimental data to determine the effects that TWF has on the performance of the Conflict Probe (CP). The null hypothesis is as follows:

> *A statistically significant Conflict Probe performance improvement is not observed through addition of the prototype enhancement of Track Weighted Descent/Climb Factor*

For this analysis, performance improvement is defined as a statistically significant reduction in False Alert Rate, no statistical difference in Late Alert Rate, and a 25th percentile of warning time that remains above the three minute threshold. All of these requirements must be true in order for it to be considered a significant improvement and to reject this null hypothesis.

The study documented here will attempt to reject this null hypothesis, therefore showing that these enhancements or parameter changes do indeed provide a significant improvement to the ERAM system. At the very least, no significant degradation must be indicated in any of the aforementioned metrics.

Table 8 shows the alert type counts for the 2 baseline runs and their respective treatment runs. A few observations can be made from it. First, there is very little difference between the MA and LA counts within each scenario (ZAU, ZDC), which suggests that no significant effects to these alerts will be evident upon statistical analysis. Second, the FA counts in the ZAU and ZDC runs are reduced by 12 and 18, respectively, from the baseline runs. This is a small change when compared to the total number of FAs in each scenario, but it a beneficial reduction and may be statistically significant. These two observations together suggest that TWF does not have a large effect on the CP.

Table 8. Alert Counts

Run	ARTCC	TWF	VA	FA	MA	LA	WT
BL1	ZAU	Off	190	1454	4	6	322
TWF1	ZAU	On	190	1442	4	6	316
BL2	ZDC	Off	203	1422	1	5	354
TWF2	ZDC	On	204	1404	0	5	351

Statistically, in order to determine whether the counts from the baseline and treatment scenarios differ in terms of the conflict probe, a chi-squared test is performed. The test statistic is defined generically as follows:

$$\chi^2 = \frac{(n_{VE} - n_{EV})^2}{(n_{VE} + n_{EV})}$$

where

χ^2 is the chi-squared test statistic

n_{VE} is the number of events of evaluation code type A in Baseline and changing to type B in Treatment

n_{EV} is the number of events of evaluation code inverse of those used in n_{VE} (type A in Treatment and type B in Baseline)

The resulting test statistic can be expressed as a probability or P-value by assuming a chi-squared distribution [Agresti, 2002]. Results of the chi-squared test are shown in Table 10 and Table 12. NS indicates a result of "not significant" and suggests that the counts in the two scenarios are not statistically different from one another. In conclusion, the results indicate there is no evidence to suggest that the TWF prototype enhancement had any significant effect on overall CP performance. However, this does not mean that individual cases do not exist, nor does this mean that there aren't differences in alerts when broken down into finer categories.

The following tables (Table 9, Table 11) detail the breakdown of alerts into various categories and indicate how the alerts associate with a given encounter vary between baseline and treatment scenarios. Labels beginning with "SAME" represent alerts that were categorized identically in both scenarios, while any other label indicates alerts that were called differently between the two scenarios. "DISCARD" in this context represents alerts that were associated with clearances (controller actions) that eventually eliminated the alert and were then excused. This breakdown and statistical testing follows. Only Valid, Missed, and False alerts are of relevance to this analysis.

Table 9. Comparison of alerts in ZAU baseline and treatment scenarios.

Evaluation Code	Description	N
FA_DISCARD	The event is a false alert in Baseline and is discarded in Treatment.	22
DISCARD_FA	The event is a discard in Baseline and is a false alert in Treatment.	20
FA_NC	The event is a false alert in Baseline and is not predicted in Treatment.	37
NC_FA	The event is not predicted in Baseline and is a false alert in Treatment.	27
SAME_FA	The event is a false alert in both Baseline and Treatment.	1395
SAME_MA	The event is a missed alert in both Baseline and Treatment.	10
SAME_VA	The event is a valid alert in both Baseline and Treatment.	190

Table 10. Significance testing of ZAU scenarios.

Comparison	Chi-Square Value	p value	Significance
DISCARD_FA to FA_DISCARD	0.095	0.758	NS
FA_NC to NC_FA	1.563	0.211	NS

For the ZAU scenario, Table 9 and Table 10 suggest that the prototype enhancement had a very slight effect on CP performance. There was a reduction in the number of false alerts (FA_NC to NC_FA), but this was not statistically significant. There was no change in Valid or Missed alerts between the two scenarios.

Table 11. Comparison of alerts in ZDC baseline and treatment scenarios.

Evaluation Code	Description	N
FA_DISCARD	The event is a false alert in Baseline and is discarded in Treatment.	16
DISCARD_FA	The event is a discard in Baseline and is a false alert in Treatment.	16
FA_NC	The event is a false alert in Baseline and is not predicted in Treatment.	43
NC_FA	The event is not predicted in Baseline and is a false alert in Treatment.	25
MA_VA	The event is a missed alert in Baseline and is a valid alert in Treatment.	1
VA_MA	The event is a valid alert in Baseline and is a missed alert in Treatment.	0
SAME_FA	The event is a false alert in both Baseline and Treatment.	1363
SAME_MA	The event is a missed alert in both Baseline and Treatment.	5
SAME_VA	The event is a valid alert in both Baseline and Treatment.	203

Table 12. Significance testing of ZDC scenarios.

Comparison	Chi-Square Value	p value	Significance
DISCARD_FA to FA_DISCARD	0	1	NS
FA_NC to NC_FA	4.764	0.029	p < .05
MA_VA to VA_MA	1	0.317	NS

For the ZDC scenario, Table 11 and Table 12 suggest that the prototype enhancement had an effect on CP performance. There was a statistically significant reduction in the number of false alerts (FA_NC to NC_FA) between the two scenarios. In addition, one MA became a VA due to the prototype enhancement, but this was not statistically significant.

3.3 Restriction Analysis

While the scope of this experiment did not allow for a detailed analysis of restrictions, their validity, and the frequency at which restrictions are actually met by the track of each flight, a cursory analysis was performed. For every "IN_EFFECT" restriction listed on a flight's plan trajectory, a latitude/longitude coordinate was obtained. This lat/long pair represents the point at which the flight is predicted to intersect a given restriction. By using these lat/long pairs and projecting them onto the actual track of each flight, the nearest point of approach of each flight to each restriction is determined. Since only a segment of the track data for each flight is available, it is necessary to filter out any cases where the nearest point of approach is determined to occur prior to or beyond the first or last track point, respectively. This process is demonstrated in **Error! Reference source not found.**.

Error! Reference source not found. depicts a flight moving from left to right. Each black sphere epresents a track point and each red sphere represents a restriction. Assume that the entire available track of the flight is presented for this example. Projections from the restrictions onto the track are required to be orthogonal to the track. In this example, the projection for restriction A (green square) falls outside the bounds of the available track data, so restriction A is not included in any analysis. The projections for restrictions B and C fall on track segments that are not beyond the bounds of the available track data and are included in any analysis. This process is not without noise, hence leeway is given when determining whether a restriction was hit or missed, described below.

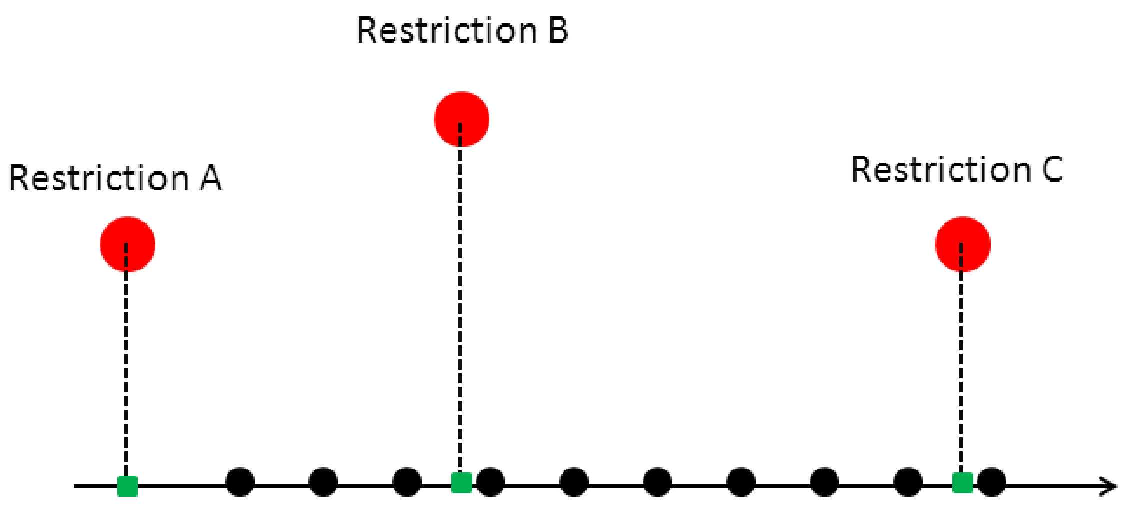

Figure 1. Projection of restrictions onto a sample track.

The analysis for the restriction data is, as previously indicated, quite simple. If a restriction is still "IN_EFFECT" when the track of a flight reached its point of nearest approach, the projection from restriction onto the track is valid as depicted in Figure 1, and the track point is within parameter horizontal distance and vertical distance of the lat/long pair for a restriction, then the flight is considered to have met the restriction. Otherwise, the flight is considered to have missed the restriction. The results of this analysis are shown in Table 13. The first set of data on the left is the result of hit/miss filtering using a horizontal threshold of 1 nm and a vertical threshold of ±300 ft. while the second set of data on the right is filtered using a horizontal threshold of 2.5 nm and ±500 ft. As depicted, restrictions are hit (within both the Horizontal and Vertical thresholds) at a rate of 53% and 56% in the ZAU and ZDC scenarios, respectively, when using the first set of thresholds. These thresholds were expanded to see if loosening the constraints would have a significant effect on hit rate. The second, looser set of parameters resulted in no improvement to hit rate for ZAU and a modest 7% improvement for ZDC. These results suggest, but do not confirm, that it is inappropriate to assume that restrictions applied to individual flight plans are always hit and should always be taken into account when trajectories are built.

Table 13. Filtered counts of hit and missed restrictions.

		H: 1.0 nm, V: 300 ft.			H: 2.5 nm, V: 500 ft.			
Center	Dimension	# Missed	# Hit	% Hit	# Missed	# Hit	% Hit	% Change
ZAU	Horizontal	224	869	80	126	967	88	+8
ZAU	Vertical	479	614	56	457	636	58	+2
ZAU	Both	574	519	53	518	575	53	+0
ZDC	Horizontal	704	2584	79	466	2822	86	+7
ZDC	Vertical	966	2322	71	900	2388	73	+2
ZDC	Both	1441	1847	56	1227	2061	63	+7

3.4 Flight Examples

3.4.1 Flight Example 1 - Improvement of Trajectory Accuracy

Example 1 (Figure 2) depicts the difference in vertical trajectory accuracy for a flight, a Boeing 737 (B737) flying level at FL330 on approach into Chicago O'Hare (ORD). Two restrictions are indicated on the plan trajectory: SOP__ORD_52/51_A_240 at FL240 and LOA_C90__ORD_51/ORD_J_VIA_BENKY_110 at FL110. The small dots represent the flight's track data in both scenarios. The aircraft trajectories, depicted as wireframes, are built at 63428 sec in both scenarios which is 90 seconds prior to the position depicted in Figure 21. Both trajectories meet the two restrictions at the bottom of each descent segment (larger blue spheres). The trajectory built in the baseline scenario (blue, no prototype enhancement) is limited in modeling the descent segments based on the track weighted factor calculated during the initial descent segment at the far left of the image. The track weighted factor for trajectory built in the treatment scenario (red) is reset after each level trajectory segment, and in this example the additional leeway in modeling descent rate produces a trajectory which is closer to the track. This provides a maximum reduction in trajectory error of 3100 ft at the point of largest deviation between the two trajectories. The trajectories are only active for 116 seconds, and the ensuring trajectories have a much lower difference in trajectory error, limiting the scenario-wide improvements to trajectory accuracy.

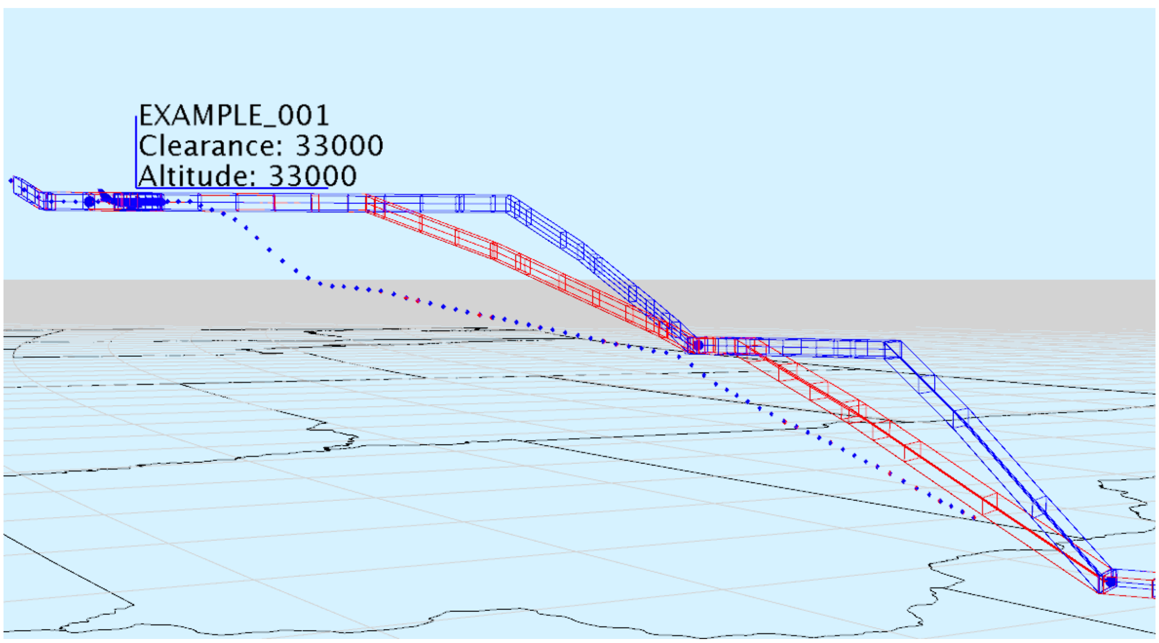

Figure 2. Trajectory improvement for a flight descending into ORD.

3.4.2 Flight Example 2 - Degradation of Trajectory Accuracy

Example 2 (Figure 3) depicts the difference in vertical trajectory accuracy for a flight, a Cessna Citation 5 Ultra Encore (C560) at about 35,000 ft on approach into Southeast Iowa Regional Airport (BRL). There are no restrictions listed on the plan trajectory, though for the purposes of trajectory modeling the destination airport is treated as a restriction. The small dots represent the past/future track of the flight. The two aircraft trajectories, depicted as wireframes, are both built at 55204 sec, the position depicted in Figure 3. The trajectory built in the baseline scenario (blue, no prototype enhancement) is constrained to use the track weighted factor calculated during the current descent segment at the far left of the image. The track weighted factor for the trajectory built in the treatment scenario (red) is reset after the level trajectory segment, and in this example the factor calculated using the prototype happens to result in a trajectory being built that increases vertical trajectory error as compared to the baseline trajectory. This provides a maximum increase in trajectory error of about 4200 ft at the point of largest deviation between the two trajectories. The trajectories are only active for 84 seconds, and the ensuring trajectories have a smaller difference in trajectory error.

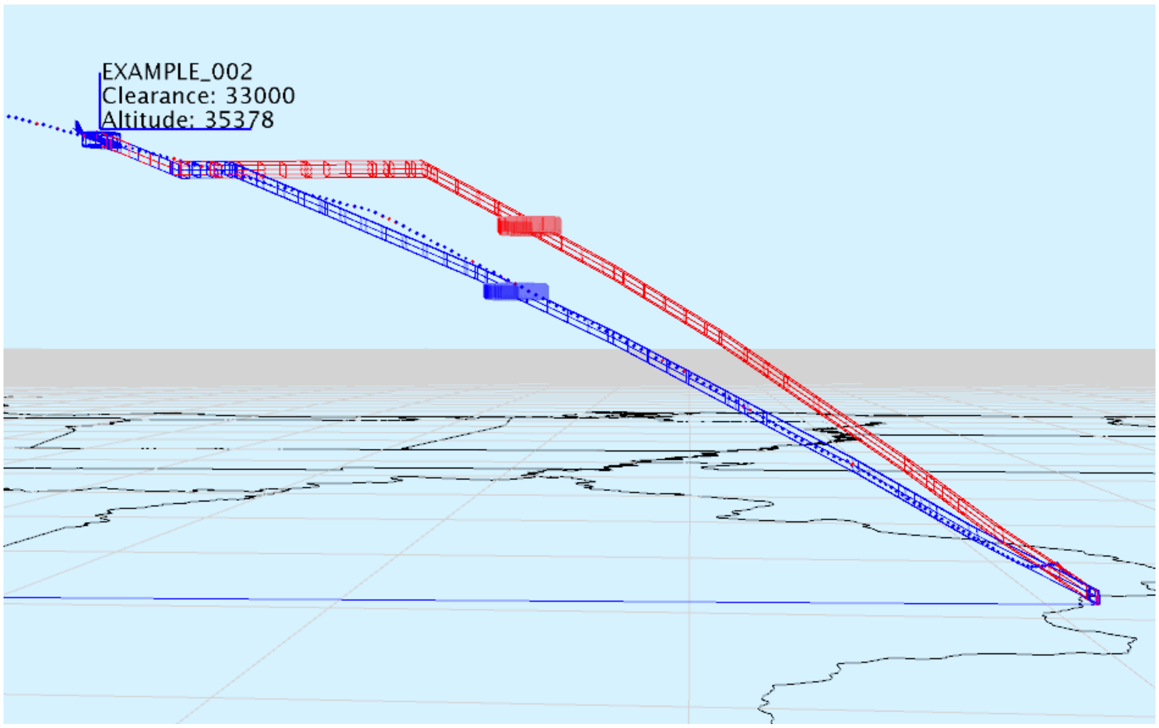

Figure 3. Trajectory degradation for a flight descending into BRL.

3.4.3 Flight Example 3 – Improvement in Conflict Probe Performance

Example 3 (Figure 4and Figure 5) depicts an encounter between EXAMPLE_003 and EXAMPLE_004 at 74461 sec. EXAMPLE_003 is a Canadair Regional Jet (CRJ7) about to descend from FL 240 to FL110 based on an interim clearance ("11000i") on approach into Chicago O'Hare (ORD) and is indicated by blue (baseline) and red (treatment) coloring. EXAMPLE_004 is a Learjet 45 (LJ45) descending from FL380 on approach into Chicago Rockford International Airport (RFD) and is indicated by black coloring. EXAMPLE_004 has just been given an interim clearance ("14000i") to descend to FL140. At this time, an alert has just been posted based on the interim clearance given to EXAMPLE_004, which removed a previously present level segment of trajectory at the point of predicted conflict. This alert, 192 seconds from currently shown position, is indicated by the slight intersection between the rectangular Conflict Probe boxes show from above in Figure 4 and from the side in Figure 5. Again, small dots represent the past/future track of the flight and wireframes represent the current aircraft trajectories for both flights. Trajectories for EXAMPLE_004 are virtually identical in the baseline and treatment scenario. However, as can be seen in the side view (Figure 5) the treatment trajectory (red) for EXAMPLE_003 has a steeper descent profile than does the baseline trajectory (blue) due to a resetting of the track weighted factor following the level segment on the left side of the image. The treatment trajectory is able to meet the restriction at the bottom of the descent leg shown (LOA_C90__ORG_51/ORD_J_VIA_BENKY_110, FL110). This is not only a better representation of the path that EXAMPLE_003 actually flew but the altered descent leg also obviates the alert called in the baseline scenario.

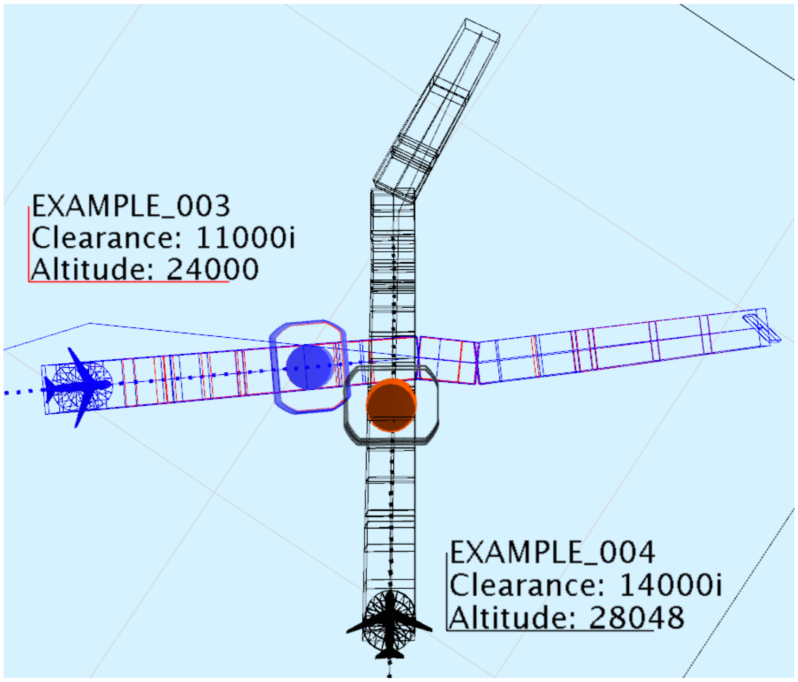

Figure 4. Top view of a crossing conflict during approach.

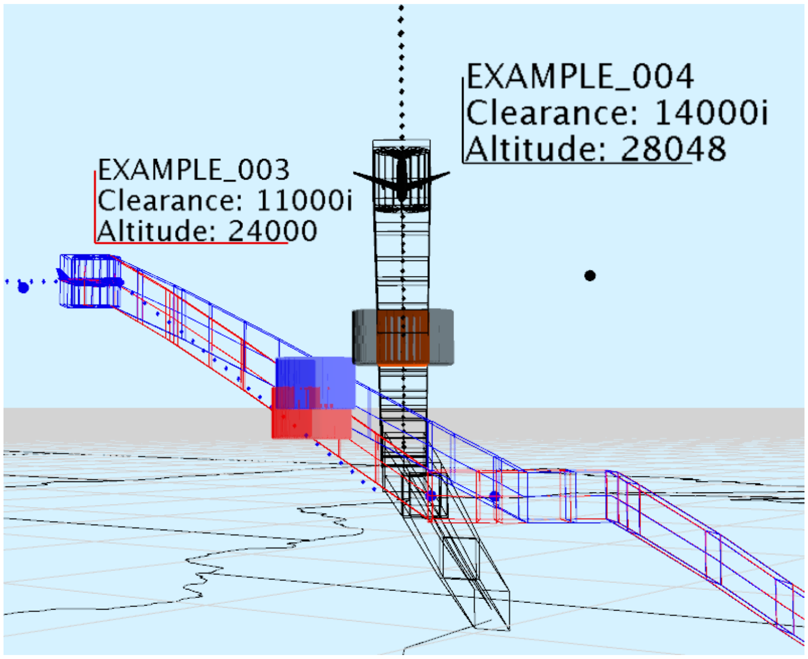

Figure 5. Side view of the same conflict during approach.

3.4.4 Flight Example 4 – Improperly Adapted Restrictions

Example 4 (Figure 6) depicts how improperly adapted restrictions can cause severely degraded trajectory accuracy through no fault of the Trajectory Modeler. Flight EXAMPLE_005, a Boeing

737-800 (B738) on approach into Chicago O'Hare (ORD), is descending from FL310 to a level segment at FL240 based on its current clearance ("14000i"). There are 3 restrictions (large blue spheres) on its plan trajectory - SOP__ORD_75/74_A_JVL30DME_240 (FL240), SOP__ORD_74/74_A_JVL15DME_110 (FL110) and LOA_C90__ORD_74/ORD_J_VIA_KRENA_100 (FL100). The small dots represent the past/future track of the flight. The aircraft trajectories, depicted as wireframes, are both built at 76072 sec, the position depicted in Figure 6. The trajectory built in the baseline scenario (blue, no prototype enhancement) is constrained to use the track weighted factor calculated during the current descent segment at the far left of the image and, because of this constraint, cannot be built in a way that meets the second restriction. The track weighted factor for the trajectory built in the treatment scenario (red) is reset after the level trajectory segment, and in this example the new factor allows the trajectory to be built in such a way that the second restriction can be met. However, as observed from the track, this is not close to the path that the flight ended up following. This is due to the fact that the second restriction is improperly adapted for the approach used by this flight, as determined in analyses performed by [Bentz and McKay, December 2012]. The TWF prototype behaves as expected given the restrictions on the flight plan; the invalid restriction itself is the cause of the degradation in trajectory accuracy. The maximum trajectory error is about 8500 ft at the point of largest deviation between the two trajectories.

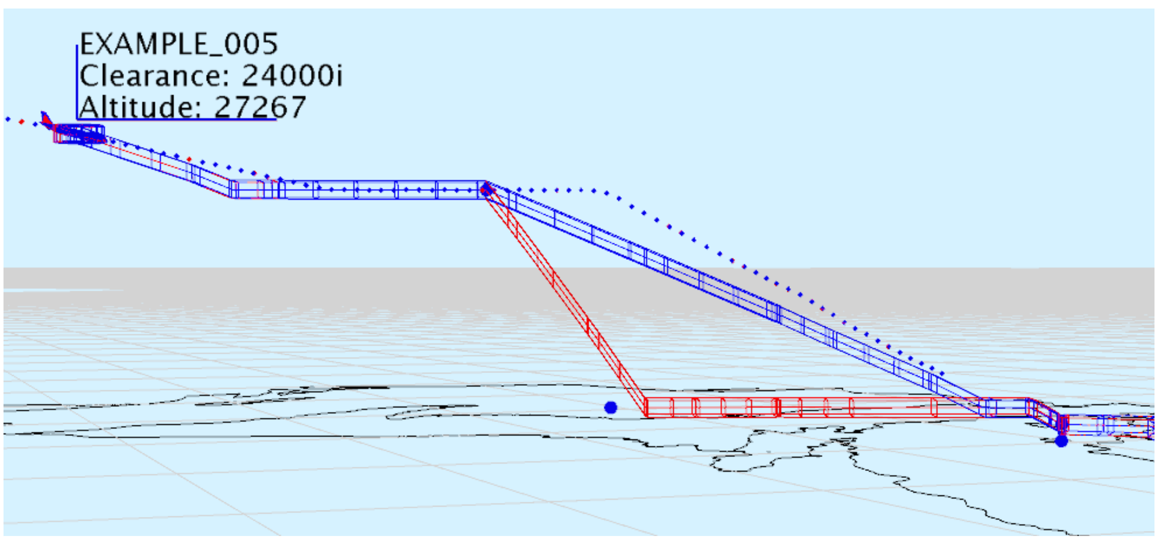

Figure 6. Improperly adapted restriction (blue sphere in the lower center of the figure).

4 Recommendations and Future Work

The TWF prototype enhancement investigated in this study has an effect on trajectory accuracy, though that effect is diluted in the current data set. A definite improvement in accuracy with respect to track on a flight by flight basis is evident, as indicated in Section 3.4.1. However, this benefit applies only for limited portions of the trajectory, a common occurrence when evaluating concepts or prototype enhancements that apply only under specific conditions. The effect of such a prototype is necessarily limited in scope and in this case results in no net improvement to vertical trajectory accuracy in the scenario as a whole. The effect on the CP is more pronounced. A trend towards reducing the amount of False Alerts in the scenario is indicated in both the ZAU and ZDC scenarios, though this trend reaches statistical significance only in the ZDC scenario.

The effectiveness of the TWF prototype, while seemingly modest, should not be considered in isolation. The Trajectory Modeler depends not only on flight plan information and recent track data, but also on the validity of the adaptation data specific to each ARTCC. If the input to the TM is of poor quality, the output will be of poor quality. Therefore, if adaptations include restrictions that have been incorrectly adapted, whether it be due to incorrect location, altitude, qualifier, or improper association with a Standard Terminal Arrival Route (STAR), the TM will often build a trajectory that is bound at least in some part by the incorrect adaptation. [Bentz and McKay, December 2012] found inaccuracies in the ERAM adaptation as pertains to restrictions (Section 3.4.4) including restrictions inappropriately associated with STARs, restrictions with incorrect qualifiers ("AT_OR_BELOW_ALTITUDE" instead of "AT_ALTITUDE") though by no means was their investigation exhaustive in this regard.

In addition, restrictions are NOT clearances, and do not specifically have to be met by aircraft. Controllers are responsible for issuing clearances so that flights meet their restrictions if desired. It's entirely possible that certain restrictions, though in effect and included in the ERAM plan trajectory, are not included in the plans or decision making processes of controllers.

The results of this study suggest that the TWF prototype is a good candidate for inclusion into ERAM. In addition to the implementation, however, follow-up work should include a thorough investigation of the validity of the restriction information currently included in ARTCC adaptations. Restrictions that are no longer current must be removed from adaptations, incorrect fields must be updated, and associations between restrictions and STARS must be verified through validation by subject matter experts. In addition, an evaluation of which restrictions are utilized by controllers and which are ignored could provide information that would be useful in tailoring the TWF prototype for specific adaptations. The improved adaptation should provide an even bigger benefit to the Trajectory Modeler.

5 List of Acronyms and Abbreviations

ANG-C41	FAA Concept Analysis Branch
ARTCC	**A**ir **R**oute **T**raffic **C**ontrol **C**enter
ATC	**A**ir **T**raffic **C**ontrol
ATO-E	Air Traffic Organization En Route Program Office
BL	FA32 **B**ase**l**ine
CP	**C**onflict **P**robe
ERAM	**E**n **R**oute **A**utomation **M**odernization
FA	**F**alse **A**lert
FAA	**F**ederal **A**viation **A**dministration
Ft	**Feet**
Horz	**Hor**i**z**ontal
LA	**L**ate **A**lert
Lat	**Lat**eral
Llh	**L**ike**lih**ood
LM	**L**ockheed **M**artin Corporation
Long	**Long**itudinal
MA	**M**issed **A**lert
NAS	**N**ational **A**irspace **S**ystem
NC	Correct **n**o-**c**all
NextGen	**Next Gen**eration Air Transportation System
nm	**N**autical **m**iles
SME	**S**ubject **M**atter **E**xpert
TM	**T**rajectory **M**odeling
VA	**V**alid **A**lert
Vert	**Vert**ical
WT	**W**arning **T**ime

References

Agresti, A. 2002. *Categorical Data Analysis*, 2nd ed., John Wiley and Sons.

Crowell, Andrew, Andrew Fabian, Christina Young, Ben Musialek, Mike Paglione. December 2011a. "Evaluation of Parameter Adjustments to the En Route Automation Modernization's Conflict Probe," FAA Technical Note, DOT/FAA/TC-TN12/2.

Crowell, Andrew, Andrew Fabian, Christina Young, Ben Musialek, Mike Paglione. December 2011b. "Evaluation of Prototype Enhancements to the En Route Automation Modernization's Conflict Probe," FAA Technical Note, DOT/FAA/TC-TN12/3.

McKay, Edward, Ph. D. February 2011. "ERAM Trajectory Lateral Modeling Enhancements, Prototype Effort," Lockheed Martin Report, Separation Management Modern Procedures.

Paglione, Mike M., Robert D. Oaks, Karl D. Bilimoria. November 2003. "Methodology for Generating Conflict Scenarios by Time Shifting Recorded Traffic Data," Proceedings of the American Institute of Aeronautics and Astronautics (AIAA) Technology, Integration, and Operations (ATIO) Technical Forum.