Trajectory Based Operations Conflict Resolution Advisories: Fast-Time Simulation Study Investigating Benefits from Improved Entry of Controller Intent

Christina M. Young[1], Andrew J. Fabian[2], and Mike M. Paglione[3]
Federal Aviation Administration William J. Hughes Technical Center, Atlantic City, NJ, 08405

This paper documents a fast-time simulation study for the NextGen Conflict Resolution Advisories (CRA) project that investigates the benefits of improved controller intent entry. CRA facilitates the entry of full 2-part amendments, and the use of these "closed clearances" will increase the level of future controller intent that is known to the ground automation system. The study employs experimental design techniques to plan and synthesize the results of 50 simulation runs that use the CRA prototype software in an En Route Automation Modernization (ERAM)-like modeling environment. Several factors are examined, including forecasted traffic demand from years 2018 and 2025, five airspace centers, and a parameter to reflect the level of intent entry to the ground automation system. These years are chosen because they represent the mid- and far-term time frame for NextGen, and five centers are selected to reflect the breadth of traffic characteristics in the National Airspace System. Metrics of interest in this study reflect trajectory prediction accuracy and conflict probe alert performance, both of which the FAA's Concept Analysis Branch (ANG-C41) Conflict Probe Assessment Team has considerable experience in evaluating. Analysis results on the effects of intent entry are presented in detail.

I. Introduction

AN advanced separation management function called Conflict Resolution Advisories (CRA) is a crucial piece of the planned implementation of the Federal Aviation Administration's (FAA) Next Generation Air Transportation System (NextGen). CRA belongs to the NextGen Trajectory Based Operations (TBO) Solution Set. TBO is an integral part of NextGen and represents a paradigm shift from today's predominantly tactical air traffic control toward strategic trajectory-based air traffic control that utilizes an unambiguous path in space and time. CRA is an advanced decision support tool (DST) designed to aid air traffic controllers in maintaining safe separation of air traffic and formulating efficient resolution maneuvers. It is ground-based and will be implemented in the En Route Automation Modernization (ERAM) system. CRA uses ERAM's conflict probe algorithm to detect potential separation violations, and provides a rank-ordered listing of potential conflict resolution maneuvers. The resolutions are presented to the controller via advanced menus accessible from the flight data block on the radar console. CRA allows the controller to easily insert their selected resolution into the ground automation system, and is expected to improve operational efficiencies as well as increase the use of "closed-loop" clearances where future controller intent is fully known to the ground automation system.

In the current system, a controller's conflict resolutions are frequently issued via voice and are often unknown to the ground automation system. In addition, the ground automation may be furnished with the first part of a maneuver but lack information about future components of the controller's intended maneuver. This benefit analysis study compares the current procedures for resolving aircraft to aircraft conflicts with procedures envisioned for future use. These future procedures will make it easier for controllers to issue clearances that contribute to a "closed-loop" system in which the ground automation system is provided with future intent and can in turn generate more accurate conflict probe results and recommended resolutions. This is contrasted to "open-loop" clearances issued today that may or may not be entered into the automation, and even when entered, do not include later maneuver components.

---

[1] Industrial Engineer, Concept Analysis Branch, WJHTC ANG-C41 300-2-K35.
[2] Computer Scientist, Concept Analysis Branch, WJHTC ANG-C41 300-2-K35, AIAA Member.
[3] Manager, Concept Analysis Branch, WJHTC ANG-C41 300-2-K35, Senior AIAA Member.

A key benefit of CRA is that it facilitates the entry of full 2-part amendments (e.g., including the second part of a 2-leg lateral maneuver or next altitude transition point). The authors hypothesize that improved intent entry leads to improved performance of the ground automation with respect to trajectory modeling and conflict alert generation. This paper documents a study testing this hypothesis in support of the cost benefit case for CRA. The second major fast-time simulation study for the CRA project, it uses the Java En Route Development Initiative (JEDI) modeling environment and Problem Analysis Resolution and Ranking (PARR) software. PARR is the CRA prototype software developed by the MITRE Corporation's Center for Advanced Aviation System Development and has been assessed in several experiments with Certified Professional Controllers (CPC) from across the country. An FAA technical note[12] documents the entire study, and the project plan is presented in Ref. 11.

A. Benefit from Improved Intent Entry

The study uses a fast-time simulation tool to investigate the benefit identified in Ref. 3 as B2: Reduced maneuvering due to improved intent entry, one of seven anticipated benefits from implementing CRA. The benefit definition is expanded to include improvements in trajectory modeling and conflict probe alert performance. The benefit from improved intent entry is associated with the use of 2-part maneuvers. CRA menus support the entry of 2-part step climbs and descents and 2-leg lateral maneuvers, and it is anticipated that the introduction of CRA will increase the frequency and accuracy of controller intent entry for these types of resolutions.

When controllers issue off-route headings without amending an aircraft's flight plan, this leaves the automation with no information on the controller's intent to return the aircraft to route; the trajectory modeler must assume a future maneuver turn point[5]. Similarly, controllers may issue temporary altitudes during climbs or descents and even though the temporary altitude may be entered into the automation, the intention of the controller regarding planned resumption of the climb or descent is not. As a result, the expected dwell time at the temporary altitude must be assumed by the trajectory modeler. In cases such as these where future intent is not entered, the presumed flight path in the ground automation has limited accuracy. The trajectories generated without the benefit of correct intent have increased potential for prediction errors, false and missed alerts, alert instability, and increased controller workload.

CRA will reduce the entry of open clearances such as temporary altitudes and off-route headings in favor of full 2-part clearance entry. When a 2-part maneuver is selected and issued to an aircraft from a CRA menu, the entire maneuver is included as a change to the known intended flight path in the ground automation. This updated intent information is incorporated in trajectory predictions used for conflict detection. This "improved intent entry" to the ground automation is anticipated to have a positive impact by reducing trajectory modeling error and the number of re-conformances, as well as improving the performance of conflict probe alerts in terms of false alerts, late alerts, and other qualities that affect controller workload.

A qualitative assessment of the benefit from CRA related to improved intent entry is detailed in a report by Kuo and Idris[7]. The report notes that "the un-ambiguous identification of aircraft intent is essential for accurate trajectory predictions, thereby allowing accurate and reliable alerting decisions at the conflict detection and resolution stages." To support this idea, Kuo and Idris present various air traffic scenarios to demonstrate how improved intent entry in CRA can result in safety benefits. This is accomplished by constructing benefit mechanisms that connect intent entry to operational errors. The simulation study provides flight data to illustrate these safety mechanisms. However, it also provides output data that will quantify the potential improvement to the trajectory and conflict probe predictions resulting from the capture of the additional intent that the CRA tool provides to the ERAM automation.

II. Study Approach

The objective of this study is to quantify the benefit to the ground automation when it is provided with the complete intended flight path for an entire maneuver when issuing 2-part maneuvers and to test the significance of any impact. To this end, the simulation and analysis is designed to determine if the automation's performance improves or degrades with increased entry of controller intent. Specific levels of intent entry are modeled by randomly withholding (to the desired degree) full amendment clearance information from the ground automation; this methodology is detailed in Section II.B. The null hypothesis to test the impact of increased levels of intent entry in this study is stated as follows:

> Regardless of en route air traffic control center and future forecasted traffic level, increasing the percent of amendment clearances provided to the ground automation (i.e. intent level) does not yield improved performance metrics.

If the null hypothesis can be rejected with a high level of confidence, this strongly implies that there is a positive impact from increasing intent levels, as stated by the alternative hypothesis:

> Increasing the percent of amendment clearances provided to the ground automation does improve the performance, as indicated by the same metrics, at different en route centers and forecasted traffic levels.

The metrics used in the analysis to capture performance are related to trajectory and conflict probe accuracy and are discussed in detail in Section II.C. The effect of the intent level was quantified using controlled experimentation techniques based on experimental design principles. The data required to assess these benefits was gathered using a fast-time simulation of the NAS. The NAS simulation was considered as a process, with input including an air traffic scenario and various controllable and uncontrollable factors, and with output consisting of the trajectory and conflict probe performance. Controllable factors are the year, the airspace, and the intent level, described in the following section.

The overall approach used in the study is as follows. The controllable factors are combined as indicated in the experimental design to produce simulation runs with various levels of each factor. Each experimental run uses a scenario file that contains one flight plan for every aircraft in that scenario. This flight plan file is passed to the fast-time simulation along with a specified intent level and other simulation-related settings. The output data from the simulation contains track data, clearance amendments, trajectories, and conflict probe alerts and is assumed to represent how the automation system would have behaved with that particular set of flight plans and the specified level of intent entry. Finally, this data is analyzed using specialized tools in order to evaluate the impact of the various factors.

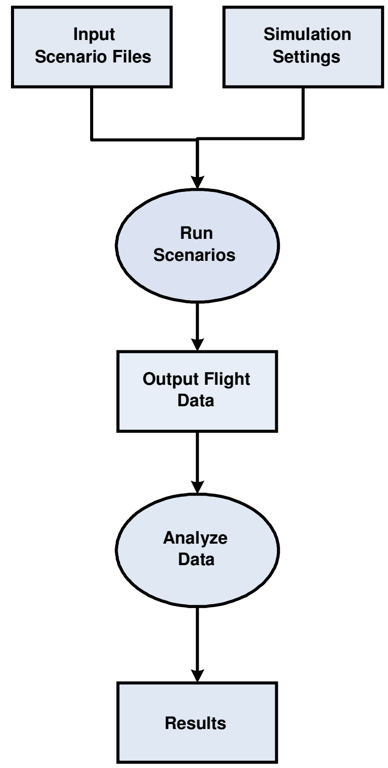

Figure 1. Study Process Flow.

Figure 1 depicts the study process. The data, represented by rectangles, consists of input flight scenario files, simulation settings, output data, and analysis results. The processes, represented by ellipses, are the fast-time simulation and analysis of data.

A. Experimental Factors and Design

Three controllable factors are considered in this study. The level of intent entry to the ground automation is the main factor of interest. Traffic density is a second experimental factor, which is represented by the year for which a traffic scenario is forecast. Finally, differences in air traffic characteristics between Air Route Traffic Control Centers (ARTCCs) may affect the performance of the ground automation, so several different centers are selected to determine how the effect from CRA may vary. The three controllable factors- percent of clearances fully entered (intent parameter), year, and ARTCC -are discussed in the following sub-sections. The combination of these factors in an experimental design is discussed in the last sub-section.

1. Intent Entry

As mentioned in the previous section, a reduction in the intent that is available to the ground automation is modeled by missing or incomplete amendment clearances. The experimental factor that represents the level of intent entry in the simulation is the percent of 2-part clearances that are fully entered, where the largest percent reflects the highest level of intent entry. Further information about the implementation of this factor is given in Section II.B. The model presented in this study allows estimation of the effect at any percent of full 2-part clearance intent entry.

Five levels of this factor will be used in this experiment, making it possible to model intent as a continuous factor. The levels of intent entry to the ground automation system are listed below, along with the 2-letter codes that are used to identify the intent parameter level in each run. Here, "clearances" refers to a full 2-part resolution maneuver.

- Full (FL), 100% of clearances entered
- High (HI), 75% of clearances entered
- Medium (MD), 50% of clearances entered
- Low (LO), 25% of clearances entered
- None (NN), no clearances entered

2. Traffic Density

Increasing levels of traffic density are simulated by using forecast traffic scenarios. The air traffic scenarios used in this study were flight plan files based on the AJG Forecast Schedules, derived from 2010 traffic levels. This study used two 24-hour scenarios: the AJG 2018 Forecast Schedule and the AJG 2025 Forecast Schedule. These years are chosen because they represent the mid- and far-term time frame for NextGen.

3. ARTCC

This study deals with conflicts identified in five ARTCCs. The five centers are selected based on operational characteristics, with the goal of selecting center facilities with different characteristics, thus representing a wide range of air traffic operations and automation performance. To aid in this selection, an analysis is performed to categorize all 20 Continental United States (CONUS) ARTCCs based on metrics for conflict probe and trajectory modeling performance and to define groups of centers with similar characteristics using statistical cluster analysis.

To provide data for the cluster analysis, a fast-time simulation is run for the 20 CONUS ARTCCs using historical track data that has been time-shifted to induce realistic conflict events. This time-shifting method has been used in previous studies[10] as a way to test the performance of trajectory modeling and conflict prediction under circumstances that closely resemble what the automation system would encounter in operation. The resulting data is analyzed to produce conflict probe and trajectory performance metrics and the cluster analysis technique is applied to the data. Cluster analysis seeks to define similar groups of entities based on their characteristics. In this analysis, Ward's clustering method is performed in JMP® using the following metrics: missed alert rate, false alert rate, average absolute cross track error, average absolute along track error, and average absolute vertical error. These metrics are discussed in more detail in Section II.C and in Ref. 2. Each of the trajectory error metrics was calculated at 5 minute and 15 minute look ahead times. Five clusters are defined as a result of the analysis, and one ARTCC is selected from each cluster to provide a wide representation of trajectory modeling and conflict probe performance. The five centers chosen for simulation are: Chicago (ZAU), Denver (ZDV), Miami (ZMA), Los Angeles (ZLA), and New York (ZNY). These centers are highlighted in Figure 2.

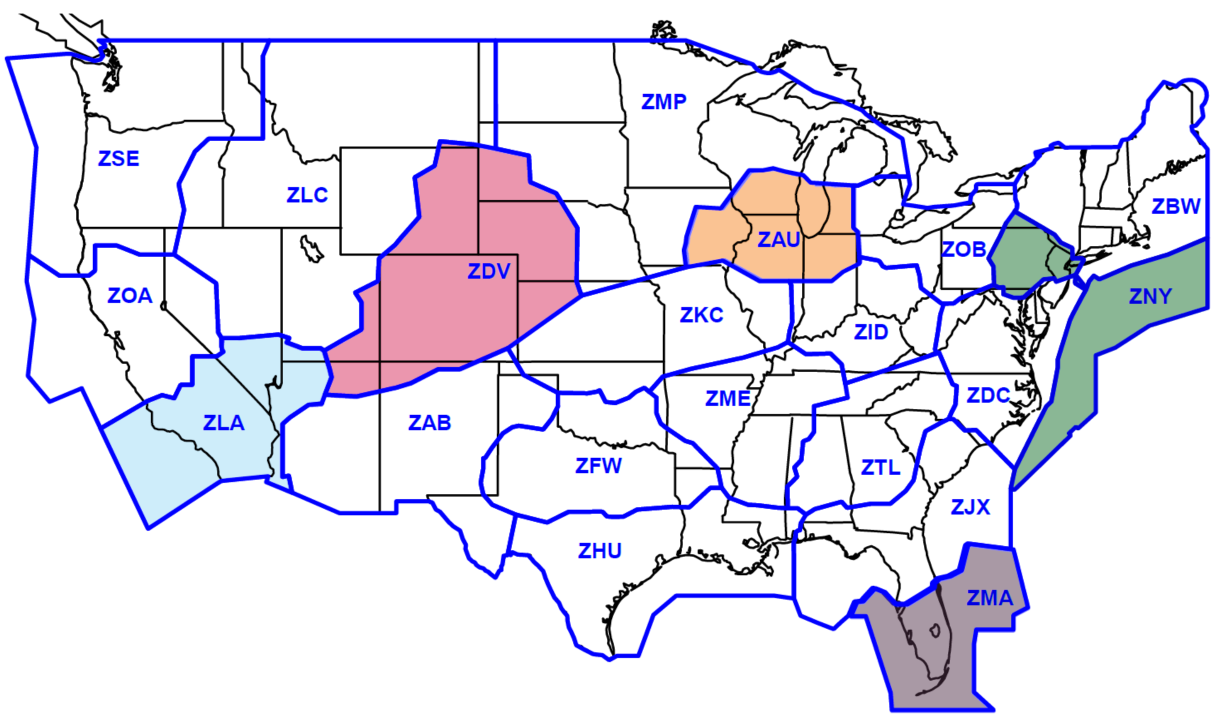

Figure 2. Air Route Traffic Control Centers.

4. Experimental Design

The levels used for each factor are: two forecast years, five ARTCCs, and five levels of intent entry. As listed in Table 1, forecast year was modeled at 2018 and 2025 traffic levels. The five ARTCCs chosen are: ZAU, ZDV, ZLA, ZMA, and ZNY. The intent parameter has 5 levels, as outlined previously. The combination of these factors at all levels produces a total of 50 possible experimental runs to study.

Table 1. Summary of Experiment Factors and Levels.

Factor	Description of Levels	Number of Levels
Intent Entry	FL, HI, MD, LO,NN	5
Traffic Density (Year)	2018 and 2025	2
ARTCC	ZAU, ZDV, ZLA, ZMA, ZNY	5
	Total Runs =	50

Since it is a fast time simulation study, the marginal cost associated with performing individual runs is relatively low. A full factorial design is executed, which includes runs that cover all possible combinations of factor levels. Full factorial designs reveal any interaction effects of the factors under study. A model will be applied to the responses from the 50 simulation runs to determine the effect of each factor.

B. Fast-Time Simulation

Fast-time simulation is used to generate track data for input flight plans, perform trajectory modeling and conflict detection, simulate CRA amendments, and output scenarios reflecting different levels of intent entry. This process is described in detail in Section 3.1 of Ref. 5 and summarized here. The fast-time simulation framework uses MITRE's Java En Route Development Initiative (JEDI) because it has trajectory modeling and conflict detection with functional performance similar to ERAM[6], which is currently being deployed and will be the operational system in all facilities.

The process starts with an input file containing one flight plan for each aircraft, from which track data is generated using a track simulator. Any predicted loss of separation between flights detected in this track data is resolved using a MITRE-CAASD problem resolution prototype named Problem Analysis Resolution and Ranking (PARR), a prototype of CRA. The resolution of notified conflicts (alerts) by PARR is invoked on a one minute cycle. For each notified conflict the highest ranked resolution is chosen and sent to the track simulator, which then simulates track data that follows the resolution.

In the full intent scenarios, the resolution amendments are also sent to the Flight Manager in JEDI for modeling. However, in a reduced intent scenario, some of this resolution information is withheld from the Flight Manager, depending on the specified intent level. This difference in information is what distinguishes the intent parameter, and it is implemented based on the factor level. For instance, at the NN or "no intent" level, the track simulator receives full two-part altitude and lateral amendments, while Flight Manager receives only interim altitudes for vertical resolution amendments, and no information at all for two-part lateral resolution amendments. At the MD or "medium intent" level, these types of resolutions have a 50% chance of being sent to the Flight Manager. Cases where Flight Manager receives missing or incomplete intent information are referred to as reduced intent amendments.

Alerts for which a resolution is sent to the track simulator are not resolved in future resolution cycles because any implemented resolution is assumed to be following a green, or conflict free, flight plan. Within JEDI a Trial Plan trajectory is used to evaluate whether a resolution path is conflict free, and the Current Plan trajectory is used by the conflict probe. For this experiment, Trial Plan conflict detection is set at a 12 minute look ahead, which effectively ensures that any proposed resolution path will be conflict free for at least 12 minutes. The Current Plan look ahead determines how far along the current flight path the conflict probe is applied and is set to 10 minutes, with resulting alert notification time between 4 and 10 minutes based on conflict likelihood.

ANG-C41 created the input flight plan scenarios for this study by using the FAA's Forecast Analysis Group (AJR-G1) Forecast Schedules. These estimate future air traffic demand and are 24-hour scenarios representing air traffic over the entire NAS, including international flights. The forecast schedules were used as a basis for air traffic scenarios representing potential flight traffic in the years 2018 and 2025. The scenario input files are generated using established tools, including the *ATOPScheduleConverter*, from the Conflict Probe Assessment Team (CPAT) within

ANG-C41. To limit each scenario to a specific ARTCC, air traffic was filtered to include only those flights traveling through some part of the ARTCC. In addition, recorded flight data was analyzed to calculate a representative distribution of aircraft equipage codes by aircraft type which was then assigned to the flight plans.

Table 2 provides a summary of the traffic counts in each of the runs for the five centers. The number of flights in each scenario output from the fast-time simulation is slightly less than the input number due to flights not reaching the center airspace before the end of the scenario and other simulation issues. These final counts are presented on the right hand side of Table 2.

Table 2. Number of Flights.

	Input Flight Plans		Simulation Output	
	2018	**2025**	**2018**	**2025**
ZAU	4931	5695	4462	5126
ZDV	3714	4326	3343	3879
ZLA	4183	4868	3795	4393
ZMA	4675	5587	4052	4790
ZNY	6532	7494	5204	5864
Total	24035	27970	20856	24052

There are 10 unique input files, which account for five different ARTCCs and two levels of traffic density. The remaining factor, intent level, was implemented as a simulation input parameter as discussed above.

C. Analysis Tools

Metrics to compare the proposed environment to the current environment are needed in order to quantify the benefit of the proposed changes. Existing tools from CPAT will be utilized in this analysis as detailed in the following sub-sections.

1. Trajectory Modeling Analysis

The output simulation data includes predictive trajectories generated by JEDI's trajectory modeling and used by the conflict probe in detection. When a reduced intent amendment is issued, the trajectory modeling system lacks the necessary information to update the cleared flight plan and build an accurate long-term trajectory. The result of this is that the flight's actual path will deviate from the known route and a new trajectory must be built. In a reduced intent scenario, the trajectory reconformance algorithm uses default parameters to estimate the turn point or altitude transition point of an aircraft in the case of a two-leg vector or step altitude maneuver, respectively. It is anticipated that in the reduced intent scenarios new trajectories will be generated more frequently, which will increase trajectory instability and degrade the performance of the conflict probe.

To quantify this change, the number of unique trajectories built by the automation for each unique aircraft identification (ACID) is recorded. Another count is made which identifies a specific subset of these trajectories. Every time an amendment is entered, a new trajectory is generated. To focus on trajectories that were built for other reasons, amendments are matched to trajectories by ACID and time (within one second) to identify cases other than when a trajectory is built following an amendment.

It is anticipated that reduced intent amendments may negatively affect the accuracy of the predicted trajectories, which also contributes to degraded conflict probe performance. Trajectory error metrics have been applied in previous studies to provide a method to measure the accuracy of trajectories in multiple dimensions with respect to actual flight position. This provides a means of quantifying the effects of improved intent. Sampling methods and definitions of these metrics are presented in Ref. 2 and summarized here briefly. The Interval Based Sampling Technique (IBST) is a method developed by ANG-C41 for evaluating trajectory accuracy. It has been previously documented in Ref. 9 and has been applied in a number of FAA studies and test programs. IBST pairs track and trajectory points to measure prediction error at various times along a flight's track and at varying look ahead times into the future.

The four basic metrics used are horizontal error, vertical error, along track error, and cross track error. Horizontal error is the time coincident difference in nautical miles (NM) between the predicted position on the trajectory and the actual position calculated from surveillance radar reports. Cross track error (NM) is the perpendicular distance between the actual position of an aircraft and its projection onto the trajectory. Along track error (NM) is the longitudinal distance along the trajectory between the same projection and the time coincident predicted position of the aircraft. The vertical error is the altitude difference in feet between the predicted trajectory position and the time coincident actual position. For further details on these definitions and how they are calculated see Ref. 9.

Following these definitions, horizontal error is unsigned while the other three are signed. For most analyses involving these metrics, it is desirable to consider absolute values because the distance from zero is of interest. Therefore the absolute value is taken before calculating average values. Finally, the average metrics comprise only relevant data points by considering error values for trajectory points where a clearance or route amendment has not been received within a specified time period (as that may have altered the trajectory) and the flight remains within control of the center.

CPAT tools are used to parse trajectory information from the JEDI output data into a database format. The information is then processed to count the total unique trajectories and unique trajectories that are not matched to a clearance amendment. Finally, the CPAT tool *TrajectorySampler* samples the trajectory error using IBST.

2. *Conflict Probe Alert Analysis*

The flight traffic in these scenarios has been processed through conflict detection, and resolution maneuvers have been implemented. As a result, there is no guaranteed way to determine from the available data whether the potential loss of separation for which an alert is generated would actually have materialized, or how close it would have been without action from air traffic control (ATC). Therefore, this study will not analyze the performance of generated alerts in terms of traditional accuracy (e.g., false alerts and missed alerts). Traditional conflict prediction metrics or a version of such will be left for future study. Instead the focus of this analysis is on the notification sets and alert-related metrics such as alert counts, duration, and predicted warning time.

CPAT tools are used to analyze the alert information output from the simulation runs to identify distinct alert notifications using rules specified in Ref. 2. The count of distinct alert notifications is compared across the scenarios and used as a response variable in the statistical model.

The duration of each alert notification is calculated as the difference between the latest alert deletion time and the earliest alert add time in a given set. It is expected that in cases of reduced intent amendments as mentioned in Section II.B, alerts will not be removed upon resolution. Since the resolution algorithm attempts to resolve conflicts every minute, alert durations greater than 1 minute are indicators of problematic alerts.

The predicted warning time provided by an alert is calculated as the notification start time subtracted from the predicted conflict start time. The distribution of warning times is analyzed across the various scenarios. A shift in the warning time distribution may indicate a degradation in conflict probe performance. Therefore the 25th percentile, or first quartile (Q1), of the alerts' predicted warning times is calculated as a metric of interest in tracking the conflict probe performance.

III. Analysis and Results

The results of the experiment are presented in three sections. First, the performance of the trajectory modeling is analyzed in terms of the accuracy of the predicted trajectory positions and the number of trajectory rebuilds that are generated during the simulation. General descriptive statistics are provided in addition to testing for statistical differences between treatment scenarios. Next, Section III.B investigates the performance of the conflict probe and specifically, the alerts generated by the simulated automation system that would have been shown to controllers interacting with the system tool. Metrics from each of these first two sections are selected as response variables in a statistical model in Section III.C, which discusses how the model is fit to the experiment data to determine the effects of the different factors and their interactions.

A. Trajectory Modeling Performance

The following two subsections provide a descriptive and inferential statistical analysis of the trajectory modeling in the simulated data. The accuracy and stability of predicted trajectories are key to the overall performance of the ERAM system and conflict probe.

1. *Counts of Trajectories*

As stated in Section II.C.1, new trajectories are built when aircraft track data do not adhere to the known route. The trajectory reconformance algorithm is forced to guess a turn point or altitude transition to rejoin the known route and this is likely to result in trajectories with poor prediction accuracy and, in turn, further trajectory rebuilds. It is expected that reduced intent scenarios will result in more frequent deviations from the known route and more frequent generation of new trajectories. The number of trajectories generated per ACID is analyzed to quantify the effect of improved intent.

The number of unique trajectories, distinguished by the trajectory build time, is recorded for each flight in a scenario. These counts are averaged over all of the flights in a scenario, and the average (per flight) values are presented in the technical note. There are clear trends in the data of count values decreasing with improved intent entry. To verify that this effect is significant, a statistical test is applied to the data. Each of the four reduced intent scenarios is considered as a treatment run and compared against the full intent scenario (FL, or 100% of clearances entered into automation) using a paired *t*-test. The same flights are present in all five scenarios for a specified Year and ARTCC, so the trajectory count for each flight in a treatment run is compared to the trajectory count for the same flight in the full intent run. The paired *t*-test examines the distribution of differences in counts between the two scenarios, and tests if the mean of the differences is statistically different from zero. One paired *t*-test is done for each of the four reduced intent scenarios in a given ARTCC and Year, for a total of 40 tests. The same conclusion is reached in all, that the trajectory counts are lower overall in the full intent scenario than in the reduced intent scenarios. This difference is statistically significant in all comparisons, with all *p*-values less than 10^{-4}. The full technical note appendices contain the difference in means and Student's paired *t*-statistic for each comparison.

It is expected that with decreasing intent levels, more clearance amendments are entered and therefore more trajectories are generated. This effect is captured by a second trajectory count metric, which focuses on the generation of additional trajectories. As explained in Section II.C.1, entered amendments are matched to generated trajectories with the same ACID and occurring within one second, and the trajectories that are not matched to a clearance are counted. This represents how many "extra" trajectories are generated in a scenario. The counts are averaged over all of the flights in a scenario, and again there are clear trends of count values decreasing with improved intent. Similar to the analysis for total trajectory count, a paired *t*-test is applied to the data to verify that the effect is significant. Comparing each reduced intent scenario to the corresponding full intent scenario, the difference for each flight is calculated. Similar to the full trajectory count analysis, 40 tests are done and the same determination of statistical significance is reached in all cases: the counts of extra trajectories in a reduced scenario are lower overall than in the full intent scenario. In these tests, the *p*-values are all less than 10^{-4} and the Student's paired *t*-statistic values are compared to the results for total trajectory counts. The results show that, even after accounting for an increase in trajectory generation from increased issuance of clearances, the effect of improved intent on trajectory generation is significant.

2. Trajectory Accuracy

To quantify the effect of improved intent entry on the accuracy of predicted trajectories, the IBST is applied to the trajectories and simulated track data. The resulting trajectory error metrics are compiled and presented here. First, the average of absolute cross track error values (denoted as AACTE) is calculated for each flight in a given scenario, taking into account all desired sampled points for that flight. The average of this value is then calculated over all flights in the scenario and presented in Table 3 to illustrate the general trend in accuracy between scenarios. The same process is used to compile average absolute along track error (AAATE) in NM, average absolute vertical error (AAVE) in feet, and average horizontal error (AHE), which is unsigned and expressed in NM. The average values by scenario for these four trajectory error metrics are presented in Table 3. The scenarios are grouped by ARTCC, Year, and Intent Level.

Table 3. Average Error Metrics by Scenario.

ARTCC	Year	Intent	Avg AACTE (NM)	Avg AAATE (NM)	Avg AAVE (ft)	Avg AHE (NM)
ZAU	2018	FL	0.030	0.143	87	0.155
		HI	0.062	0.169	97	0.205
		MD	0.106	0.203	111	0.272
		LO	0.157	0.237	127	0.343
		NN	0.208	0.244	143	0.388
	2025	FL	0.028	0.140	87	0.151
		HI	0.061	0.176	100	0.211
		MD	0.112	0.203	116	0.276
		LO	0.160	0.249	132	0.354
		NN	0.227	0.260	147	0.415

ARTCC	Year	Intent	Avg AACTE (NM)	Avg AAATE (NM)	Avg AAVE (ft)	Avg AHE (NM)
ZDV	2018	FL	0.021	0.044	13	0.049
		HI	0.046	0.063	17	0.089
		MD	0.075	0.074	19	0.124
		LO	0.112	0.093	23	0.171
		NN	0.132	0.096	28	0.190
	2025	FL	0.021	0.042	15	0.050
		HI	0.058	0.071	19	0.106
		MD	0.076	0.078	21	0.128
		LO	0.107	0.088	28	0.164
		NN	0.136	0.103	32	0.199
ZLA	2018	FL	0.026	0.088	52	0.098
		HI	0.053	0.123	61	0.145
		MD	0.105	0.161	72	0.227
		LO	0.157	0.194	79	0.283
		NN	0.227	0.235	97	0.367
	2025	FL	0.028	0.093	55	0.103
		HI	0.079	0.141	69	0.192
		MD	0.160	0.194	77	0.301
		LO	0.216	0.230	95	0.378
		NN	0.312	0.296	116	0.500
ZMA	2018	FL	0.022	0.128	74	0.124
		HI	0.039	0.155	86	0.159
		MD	0.071	0.178	103	0.210
		LO	0.097	0.200	116	0.246
		NN	0.126	0.205	132	0.273
	2025	FL	0.023	0.139	74	0.133
		HI	0.055	0.162	87	0.180
		MD	0.073	0.183	104	0.215
		LO	0.113	0.201	121	0.263
		NN	0.166	0.225	142	0.325
ZNY	2018	FL	0.036	0.234	88	0.220
		HI	0.072	0.279	102	0.286
		MD	0.090	0.308	108	0.324
		LO	0.134	0.327	127	0.379
		NN	0.169	0.348	136	0.420
	2025	FL	0.039	0.234	94	0.226
		HI	0.061	0.282	109	0.283
		MD	0.097	0.295	114	0.323
		LO	0.139	0.349	132	0.400
		NN	0.176	0.368	146	0.440

In Table 3 there are clear trends of trajectory accuracy increasing (error values decreasing) with improved intent. To verify that this effect is statistically significant, a paired *t*-test is applied to the data. Each of the four reduced intent scenarios is compared against the corresponding full intent scenario. The same flights are present in all five scenarios for a specified Year and ARTCC, and the average trajectory error for each flight in a reduced intent scenario is compared to the average error for the same flight in the full intent scenario. Four paired *t*-tests (one each for AACTE, AAATE, AAVE, and AHE) are done for each of the four reduced intent scenarios in a given ARTCC and Year, for a total of 160 tests. The same conclusion is reached, that trajectory errors are lower overall in the full intent scenario than in the reduced intent scenarios. This difference is statistically significant in all comparisons, with all *p*-values less than 10^{-3}. The technical note appendices contain the difference in means and Student's paired *t*-statistic for each comparison.

A different analysis approach is to consider the trajectory errors at a specific amount of time into the future, or look ahead time. This allows for evaluation of the trajectory predictor performance trend with look ahead time, and how this trend is affected by improved intent. The average absolute cross track error for a given look ahead time and all sampled points is calculated for 0, 300, 600, 900, and 1200 second look ahead times (every 5 minutes) and presented in Figure 3 for the set of ZAU (2018 and 2025) scenarios. The legend lists the scenarios in the same order in which they appear in the graphs.

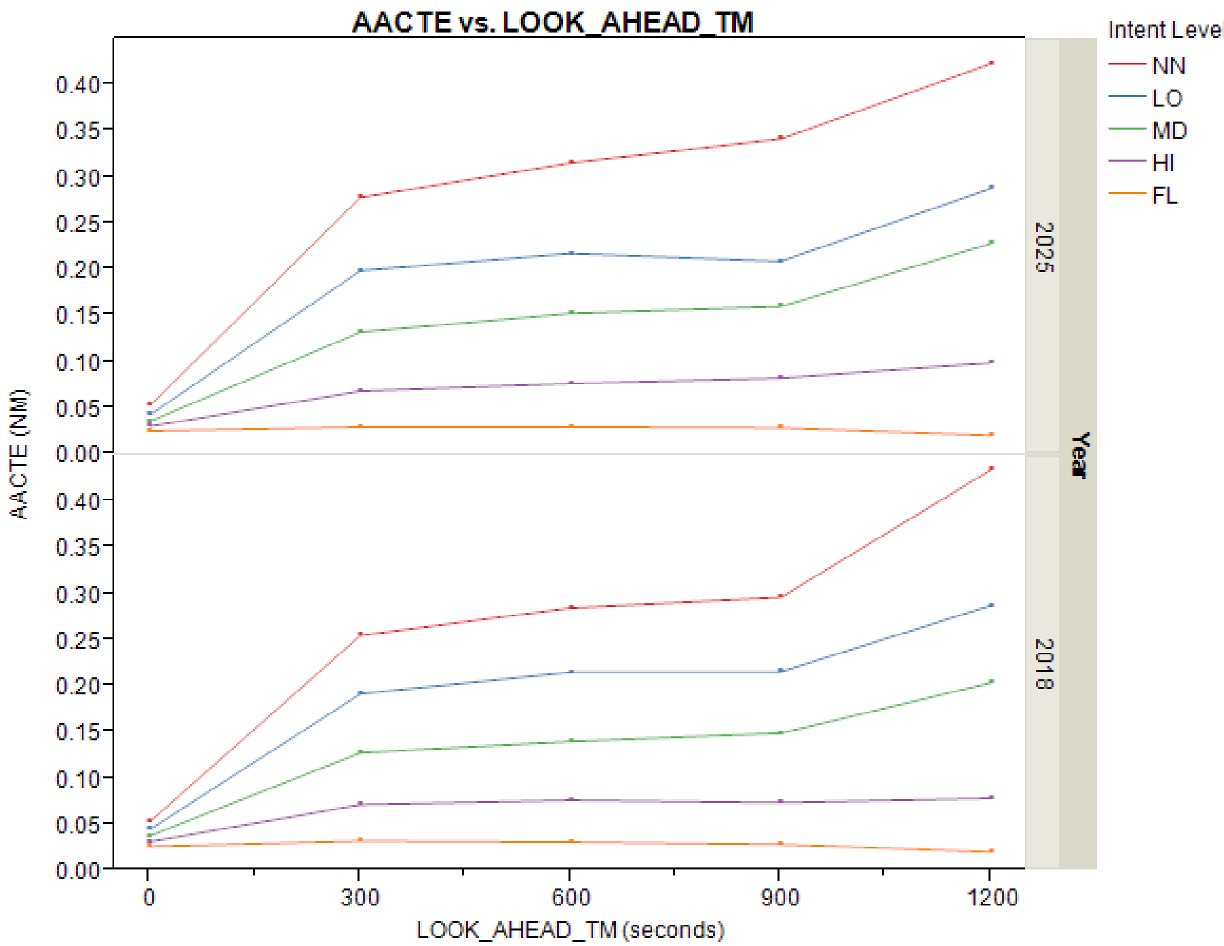

Figure 3. Trajectory Error vs. Look Ahead Time for ZAU.

From Figure 3 it is observed that an increase in the level of intent entry corresponds to a decrease in average absolute cross track error at every look ahead time for the ten scenarios shown. In addition, the benefit from increased intent entry is more pronounced at longer look ahead times. Similar graphs for vertical and along track errors are located in the technical note appendices. It is noted that the relative differences are smaller for vertical error, and there is a slight inconsistency in the pattern for along track error.

B. Conflict Probe Alert Performance

Conflict alerts generated by the automation are collected for each simulated scenario. Alert addition, modification, and deletion events are grouped into notification sets using CPAT tools with specially designed logic. These notification sets are analyzed for overall count, alert duration, and predicted warning time to demonstrate benefits from improved intent entry. The count of distinct alert notifications is presented in Figure 4.

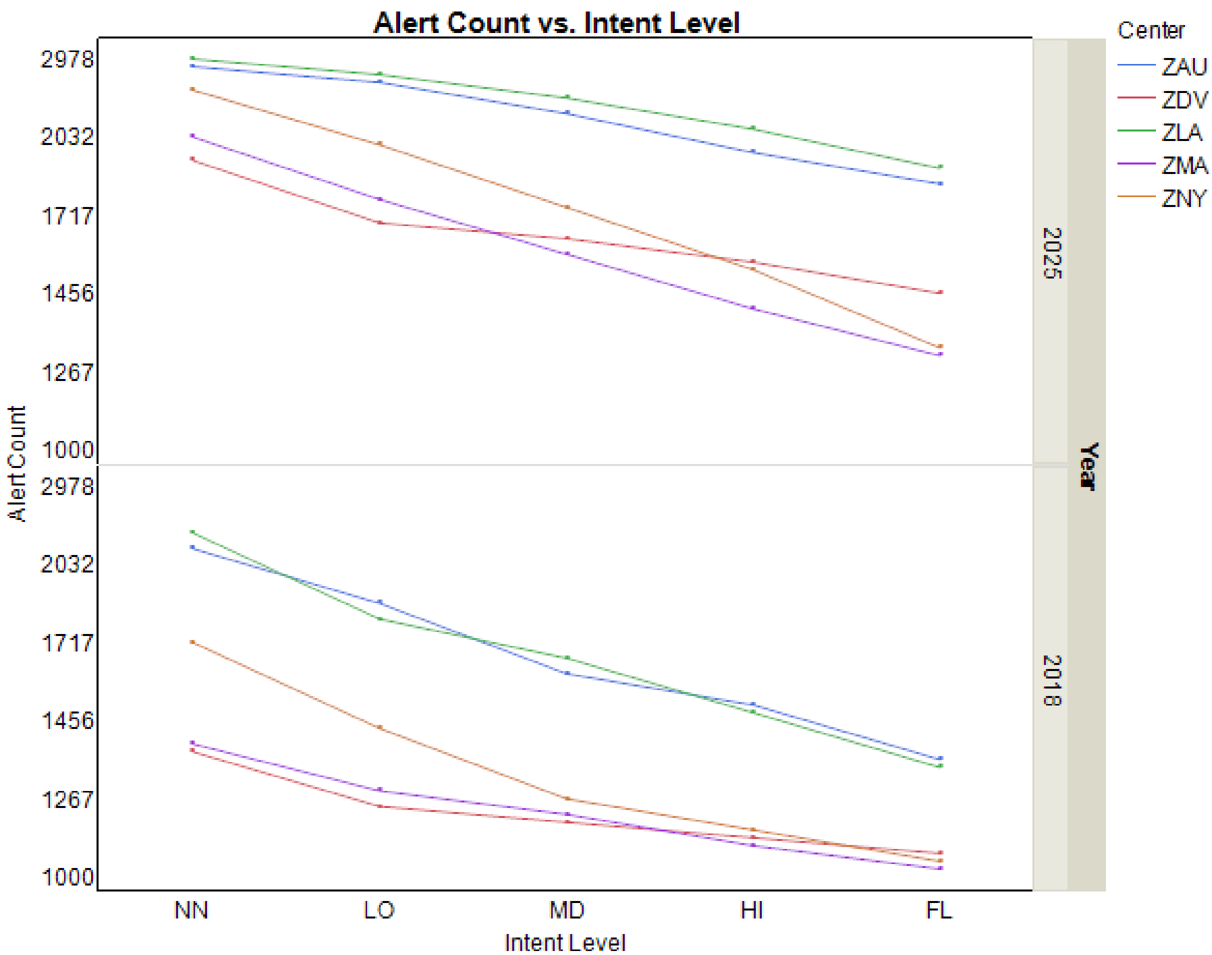

Figure 4. **Alert Count vs. Intent Level.**

From Figure 4, the trend in each center is a consistent decrease with increased intent level, regardless of traffic year. These trends differ across the various centers. In particular, ZDV exhibits slightly different behavior, although the counts do follow the general trend of decreasing with increased intent level. These types of differences may be investigated in future studies.

Due to open clearances in reduced intent scenarios, it is expected that a significant number of alerts will not be removed upon issuing an amendment, whereas with full intent, the majority are successfully resolved and the alerts deleted. Alerts with duration greater than 60 seconds are depicted in Figure 5, which provides the distribution of alert duration in one minute intervals. Since the resolution algorithm attempts to resolve alerts every minute, these can be interpreted as alerts that are not deleted at the time an amendment is entered. The percentage of problematic events over all events allows for a relative measure of performance between the scenarios. It is expected that improving the ground automation will diminish this percentage and it is hypothesized here that providing better intent information will improve the related automation functions. The metric 'percent of alerts with long duration' was computed for all scenarios and aggregate percentages are, from most intent to least intent provided: 15.6%, 21.5%, 26.3%, 28.9%, and 34.0%. These data are further detailed by partitioning the event durations into increments of one minute. Figure 5 illustrates the metric over the five intent levels for ZAU with 2018 traffic and provides evidence that as intent is increased, the percentage of problematic events decreases. The other scenarios show a similar relative pattern and are included in the technical note appendices.

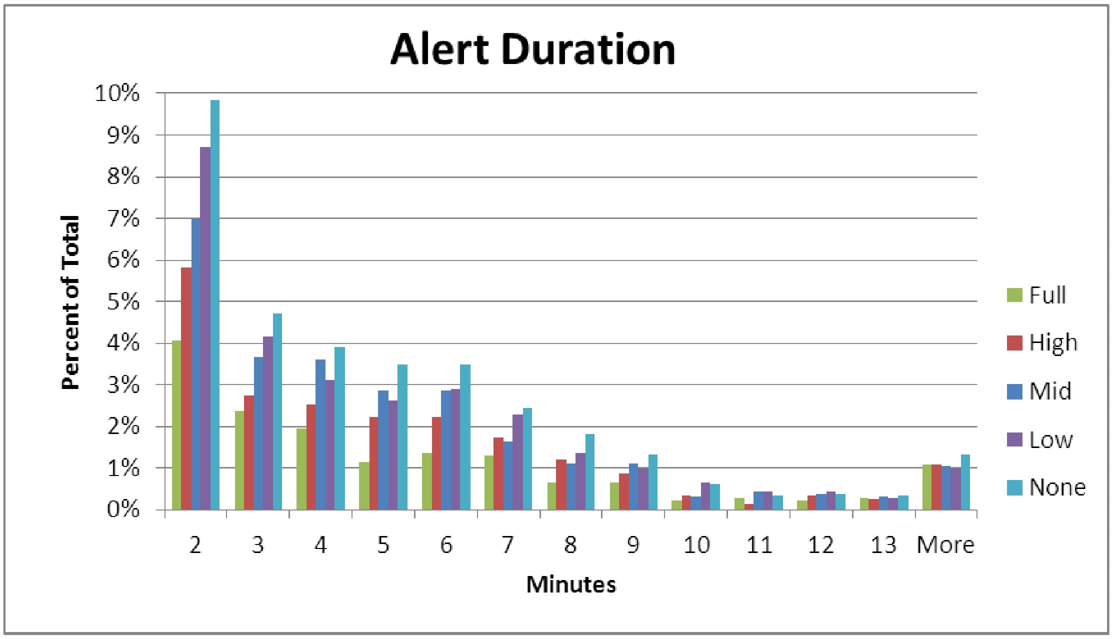

Figure 5. **Percent of Alerts with Duration Exceeding One Minute.**

Another result of inaccurate trajectories from reduced intent amendments is an increase in the frequency of late notification of alerts. One factor that contributes to this increase is a short problem detection look ahead during the first off-leg of a vector maneuver or the level segment of a step altitude maneuver. For instance, if a modeled segment length has shorter duration than the controller's true intention, potential conflicts may only be alerted after a reconformance trajectory extending the off-leg or level segment is built, and accordingly these alerts may have a relatively small warning time. The first quartile of the predicted warning time is a valuable metric to describe how the lower tail of the distribution is affected by a change.

To demonstrate how the predicted warning time distribution is affected, Table 4 presents data from ZAU, grouped by intent level and year. Count is the number of alerts generated, and Percent of Alerts with Duration > 1 min. represents the aggregated percentages from the previous histograms. Q1 of Predicted Warning Time is the 25th percentile of the warning times predicted by the alerts (predicted conflict start time – notification start time) in seconds.

Table 4: Alert Statistics for ZAU Scenarios.

ARTCC	Year	Intent	Count	Q1 of Predicted Warning Time (seconds)	Percent with Duration > 1 min.
ZAU	2018	FL	1381	291	15.6
		HI	1494	282	21.5
		MD	1635	270	26.3
		LO	1829	240	28.9
		NN	2103	199	34
	2025	FL	1812	292	14.3
		HI	1939	283	18.9
		MD	2116	273	24.1
		LO	2351	244	28.9
		NN	2713	208	33

In general, the number of alerts created increases as the intent level decreases, as is expected, though there is one anomalous value for ZDV 2018 with low intent level. Q1 of the predicted warning time represents the minimum predicted warning time for three-quarters of the data set, so decreasing this value represents a degradation of the minimum expected warning time predicted by the majority of alerts. This metric provides an alternate way of showing change in performance and supports the histograms, represented in aggregate form in the tables as Percent Alerts with Duration > 1 min. They support each other since one shows degradation in performance via a decrease in predicted warning time and the other shows this via an increasing percentage of problematic alerts. Similar tables for the other four centers are provided in the technical note.

C. Statistical Model of Results

The results of implementing the inferential statistical approach are presented here. A detailed multivariate regression model is used to fit the results of the designed experiment with the goal of determining which factors have a significant effect on the response variables and the size of these effects. Metrics from trajectory modeling and conflict alert performance are selected as response variables for the model. The first sub-section implements a statistical model and describes how the experiment data is fitted to the model, while the last discusses the findings from the model.

1. Model Implementation

Equation 1 illustrates the mathematical model for this experiment. It represents the full factorial design where all levels and factors are crossed, allowing all the interactions to be examined. This amounts to three main effects (single variables), three two-way interaction terms (double variables), and a quadratic term on the continuous variable, intent level. The constant or overall mean effect is represented as the "μ" term.

Response:

$$R_{ijk} = \mu + Y_i + A_j + I_k + Y_i \times A_j + Y_i \times I_k + A_j \times I_k + I_k \times I_k + \varepsilon_{n(ijk)}$$

Where:

$$\begin{aligned} &Y_i = \text{forecast years, } i = 1, 2 \\ &A_j = \text{ARTCC, } j = 1, 2, 3, 4, 5 \\ &I_k = \text{intent level, } k = 1, 2, 3, 4, 5 \\ &\varepsilon_{n(ijk)} = \text{random error, } n = 1, 2, \ldots \text{ for all } i, j, k \end{aligned} \tag{1}$$

The model assumes the random error $\varepsilon_{n(ijk)}$ is approximately independently normally distributed with a zero mean and that the various factors are linearly additive as illustrated in Equation 1.

Five response variables are evaluated and the same model in Equation 1 addresses all five separately. Thus, the term "R_{ijk}" can refer to any one of the response variables. These values are calculated for each of the various runs (and associated factor levels) defined in Table 1. The term "$R_{ijk,}$" then, is an estimate of the expected value for each of these five output functions. The response variables studied with this model are: average absolute horizontal trajectory error, average absolute vertical trajectory error, the count of trajectories with no associated clearance, the first quartile of predicted warning time, and the count of alerts with duration greater than one minute.

The experimental design coded in Equation 1 and presented in Section II.A.4 is a full factorial design[8], which includes all possible combinations of factor levels in the experiment. It is expensive in terms of runs required but offers several advantages, especially early in the study of a process. In this study, the quantity of runs is relatively inexpensive because a fast-time simulation model is employed. Factorial designs can be used to reveal the interaction effects of the factors under study and they are significantly more efficient than running multiple experiments for one factor at a time. The combinations of factor levels provide replications for evaluation of the individual factors, when some factors or factor combinations are removed from the experiment. The full factorial experiment is implemented, results from the 50 experimental runs are collected, and the model is fit to the calculated response data.

2. Model Findings

The fitted model is summarized graphically in Figure 6, where five leverage plots illustrate the actual and modeled values for each of the five responses. If the model could perfectly capture all the observed variation in the system, the actual measured response mean plotted on the y-axis in the figures and the coincident modeled version on the x-axis would fall on a diagonal line perfectly. The term "Rsq" in the plots is the coefficient of determination of the model.* This term provides a quantification of how well the model captures the observed variation in the system under study. For the five response variables under study in this experiment, the R^2 ranged from 0.97 to 1.00. In practical terms, this means that the model defined in Equation 1 captured from 97 to 100 percent of the variation in the actual system under study. It is clear for all five responses that the model captures the trend and a high percentage of the variation.

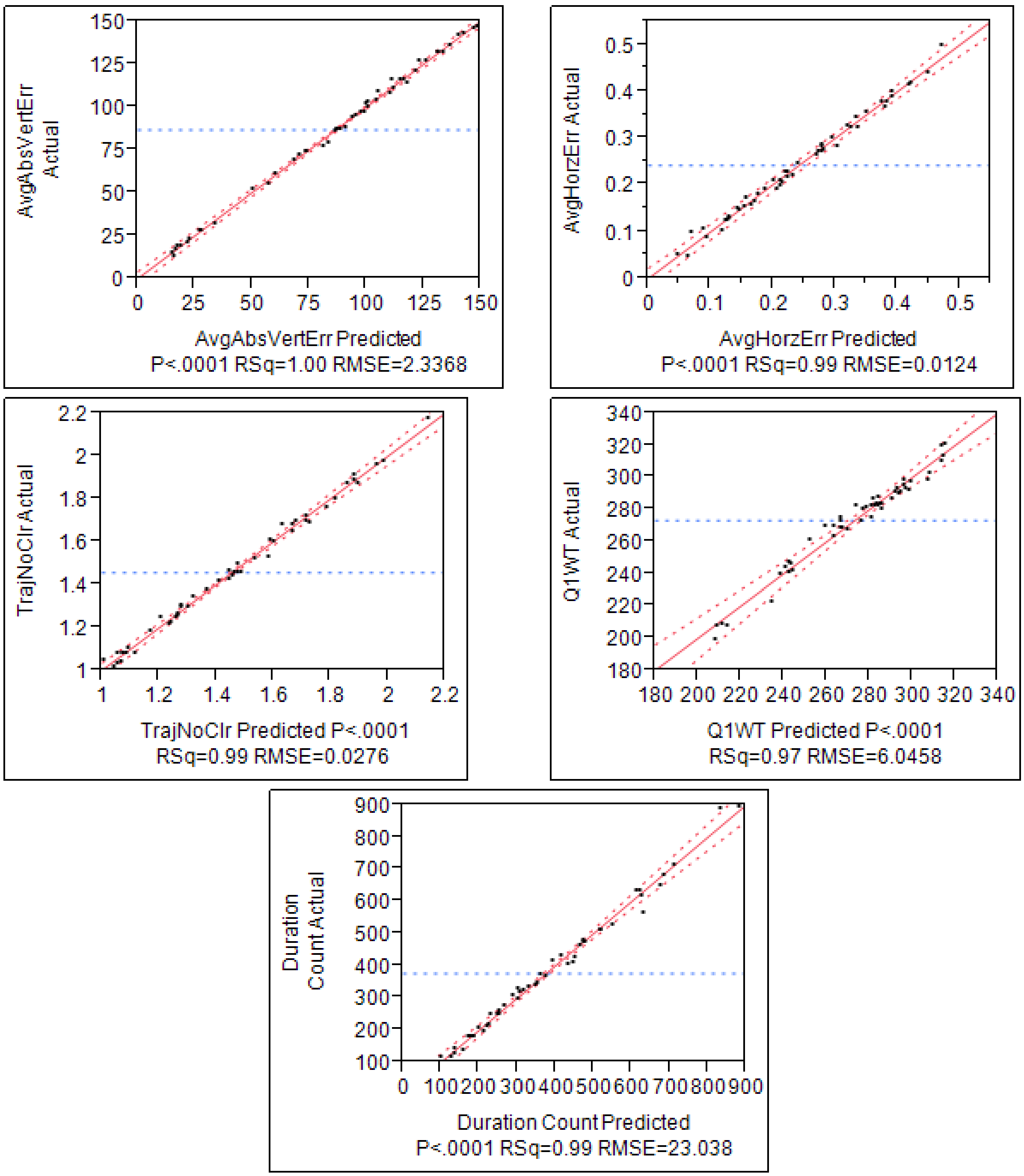

Figure 6. Leverage Plots per Response.

* From Ref. 4, the R^2 is the coefficient of determination and is equal to the ratio of the sum of squares of the model divided by the sum of squares of the total variation. The total variation equals the modeled variation plus the error in the model (estimated by calculating the difference between model and observed values).

Tables 5 through 9 list the effect tests for the various factor level combinations of the experiment. The intent level is a continuous factor that may have non-linear effects while the others all represent fixed effects. The column labeled “Source” defines the particular effect produced from the combinations of factors listed. The column labeled “DF” is the degrees of freedom for the particular factor combination. The column labeled “Sum of Squares” is calculated by summing the squared differences of the observations minus the mean. The column labeled “F Ratio” is the test statistic produced by model mean square divided by the error mean square. The column labeled “*p*-value” is the probability that the test statistic is not significant. A *p*-value that is less than 0.05 is marked by an asterisk to indicate it provides evidence that the particular factor is statistically significant.

Table 5. Model Effect Tests for Average Horizontal Error.

Source	DF	Sum of Squares	F Ratio	*p*-Value
Intent%	1	0.12746576	832.8155	<.0001*
ARTCC	4	0.11483390	187.5709	<.0001*
Traffic Year	1	0.00651133	42.5427	<.0001*
Intent%*ARTCC	4	0.02748042	44.8868	<.0001*
Intent%*Traffic Year	1	0.00242900	15.8702	0.0004*
ARTCC*Traffic Year	4	0.00760976	12.4299	<.0001*
Intent%*Intent%	1	0.00000131	0.0086	0.9267

Table 6. Model Effect Tests for Average Absolute Vertical Error.

Source	DF	Sum of Squares	F Ratio	P-Value
Intent%	1	6058.225	1109.408	<.0001*
ARTCC	4	30930.359	1416.025	<.0001*
Traffic Year	1	371.661	68.0601	<.0001*
Intent%*ARTCC	4	1755.926	80.3882	<.0001*
Intent%*Traffic Year	1	72.955	13.3599	0.0009*
ARTCC*Traffic Year	4	103.784	4.7513	0.0039*
Intent%*Intent%	1	70.119	12.8405	0.0011*

Table 7. Model Effect Tests for Response Variable TrajNoClr.

Source	DF	Sum of Squares	F Ratio	P-Value
Intent%	1	1.7607103	2317.748	<.0001*
ARTCC	4	0.0991183	32.6191	<.0001*
Traffic Year	1	0.0360800	47.4946	<.0001*
Intent%*ARTCC	4	0.1171211	38.5437	<.0001*
Intent%*Traffic Year	1	0.0158811	20.9054	<.0001*
ARTCC*Traffic Year	4	0.0112091	3.6888	0.0137*
Intent%*Intent%	1	0.0037736	4.9675	0.0328*

Table 8. Model Effect Tests for Response Variable Q1WT.

Source	DF	Sum of Squares	F Ratio	P-Value
Intent%	1	10432.901	285.4311	<.0001*
ARTCC	4	6417.190	43.8916	<.0001*
Traffic Year	1	25.205	0.6896	0.4123
Intent%*ARTCC	4	2834.403	19.3864	<.0001*
Intent%*Traffic Year	1	45.901	1.2558	0.2705
ARTCC*Traffic Year	4	145.833	0.9974	0.4228
Intent%*Intent%	1	1477.125	40.4123	<.0001*

Table 9. Model Effect Tests for Response Variable DurationCount.

Source	DF	Sum of Squares	F Ratio	P-Value
Intent%	1	480592.08	905.5296	<.0001*
ARTCC	4	109485.04	51.5728	<.0001*
Traffic Year	1	164393.78	309.7501	<.0001*
Intent%*ARTCC	4	51098.96	24.0701	<.0001*
Intent%*Traffic Year	1	31612.84	59.5648	<.0001*
ARTCC*Traffic Year	4	6461.92	3.0439	0.0306*
Intent%*Intent%	1	24182.86	45.5652	<.0001*

The results in Table 5 through Table 9 indicate that all of the main factors had a statistically significant effect according to the fitted model, with the exception of traffic year (and its possible interactions) on predicted warning time. Furthermore, the effect of intent level is found to be non-linear for all response variables, with the exception of horizontal error, for which intent has a fixed, linear effect.

The model assumes that the unattributed variation or error in the model, referred to as random error, $\varepsilon_{n(ijk)}$ in Equation 1 is approximately normally distributed. An additional validation of the model is to test the residuals for normality. In Ref. 12 these residual errors are presented for each of the five response variables in histograms overlaid with fitted normal distribution density lines, box plots, and normal probability plots for each response variable. The histograms and box plots illustrate that the distributions are fairly symmetric and centered at zero as expected if normally distributed. The normal probability plot illustrates for each response that the model errors fall along the diagonal probability line, indicating that each is at least approximately normally distributed and supporting the validity of the model.

The experimental results produce a statistical model with coefficient estimates that are summarized in Table 10. This model allows us to draw conclusions on the relationships and net effects of the various factors under study.

Table 10. Summary of Model Coefficient Estimates.

Source	Avg Horz Err (NM)	Avg Abs Vert Error (ft)	Trajectory (No Clr.) Count	Duration Count	Q1 of Warning Time (s)
Intercept	0.331	104.47	1.7937	488.27	251.476
Intent%	-0.002	-0.44	-0.0075	-3.92	0.578
ARTCC[ZAU]	0.043	29.28	0.0434	123.52	-17.47
ARTCC[ZMA]	-0.028	18.07	-0.0127	-61.88	0.48
ARTCC[ZNY]	0.096	28.10	0.0284	-45.28	-14.07
ARTCC[ZLA]	-0.006	-11.58	0.0576	9.52	3.33
ARTCC[ZDV]	-0.105	-63.86	-0.1168	-25.88	27.73
Traffic Year[25-18]	0.0228	5.45	0.0537	114.68	-1.42
(Intent%-50)*ARTCC[ZAU]	-0.000	-0.11	-0.005	-0.94	0.26
(Intent%-50)*ARTCC[ZMA]	0.001	-0.16	0.0004	0.98	-0.25
(Intent%-50)*ARTCC[ZNY]	0.000	-0.02	0.0005	0.46	0.22
(Intent%-50)*ARTCC[ZLA]	-0.001	-0.03	-0.0023	-1.23	-0.01
(Intent%-50)*ARTCC[ZDV]	0.001	0.32	0.0019	0.73	-0.21
(Intent%-50)*Traffic Year[25-18]	-0.000	-0.07	-0.0010	-1.42	0.05
ARTCC[ZAU]*Traffic Year[25-18]	-0.014	-2.24	-0.0301	-14.68	4.97
ARTCC[ZMA]*Traffic Year[25-18]	-0.002	-1.67	-0.0036	4.12	-3.48
ARTCC[ZNY]*Traffic Year[25-18]	-0.014	1.50	-0.0274	-32.88	2.02
ARTCC[ZLA]*Traffic Year[25-18]	0.048	4.98	0.0523	34.52	-4.13
ARTCC[ZDV]*Traffic Year[25-18]	-0.018	-2.57	0.0089	8.92	0.62
(Intent%-50)*(Intent%-50)	0.000	0.00	0.0000	0.02	-0.01

The results in Table 10 can be interpreted by comparing the size of the effects to the intercept, which is the mean response value over all levels. For every one percent increase in intent entered, the average horizontal error decreases by 0.002NM and average absolute vertical error by -0.44 ft. The number of trajectories per flight decreases by 0.0075 for every percent increase in intent entry, compared to the 1.7937 overall average. The count of alerts with duration greater than one minute decreases by roughly 3.92 on average per one percent increase. This is compared to the intercept value of 488 total long duration alerts. The first quartile of predicted warning time increases by 0.578 seconds for every increase of one percent intent entry. Comparing no intent entry (0%) to full intent entry (100%), therefore, the first quartile increases by roughly one minute. Likewise, the horizontal error decreases by 0.2NM, or 61%; vertical error decreases by 44 ft, or roughly 42% of the general average. The trajectory count per flight decreases by 0.7 from minimum to maximum intent entry, which is 42%. And the count of alerts with long duration decreases by 392, or 80% of the intercept value.

The JMP® commercial software tool from SAS provides an interactive model calculator called the predictor profiler that allows the examination of the effects of the various factors of the model. Figure 7 presents the predictor profile plot of the model results. The general trend is that year has a negligible effect on all responses except the count of alerts with long duration. The choice of ARTCC has the largest effect on the average of the absolute vertical errors. Increasing the intent level also displays desirable effects in the model: average horizontal and vertical error decrease, the count of trajectories with no associated clearance decreases, the count of alerts with long duration decreases, and the predicted warning time increases. These trends show that the model agrees with the hypothesis that increased intent in the system produces a positive effect.

The slopes of the plotted lines in Figure 7 indicate the magnitude and direction of each factor's effect on the model. The curvature in the intent factor indicates its non-linear effects which are stronger at low intent levels. The y-axis plots the response variable estimates from the model and the decimal numbers on each y-axis represents the modeled response variable at the levels specified in the figures. The setting chosen for the profiler graphics are the ZDV ARTCC, at full (100%) intent for 2018. ZDV is chosen because it demonstrates the greatest positive benefits.

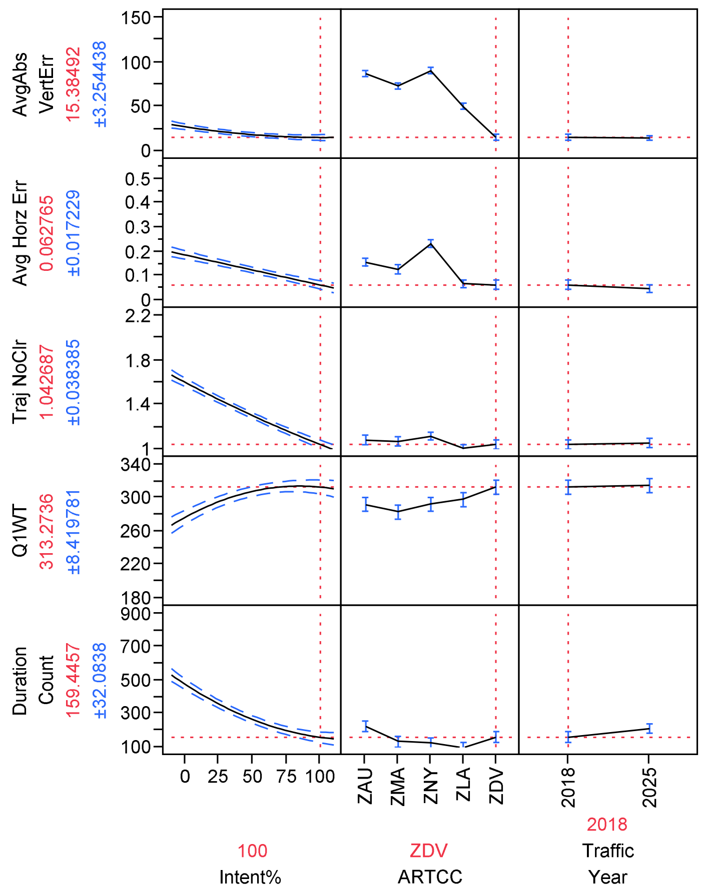

Figure 7. Predictor Profiler for ZDV 2018.

IV. Conclusion

CRA is envisioned as an advanced decision support tool for air traffic control deployed within the NextGen TBO initiatives. It predicts future conflicts between aircraft and proposes ranked resolution options, supporting increased use of "closed-loop" clearances by facilitating the entry of controller-selected resolutions. Contrasted with the methods used in today's operations, where amendments are typically issued via voice and are often unknown to the ground automation system, this improved entry of amendments expands the information available to the automation regarding future intent. It is hypothesized that this additional information will increase the accuracy of ground-based trajectory modeling and improve conflict probe alert performance.

This is one of a series of studies to estimate a number of potential benefits of CRA. The objective of this study is to investigate the benefits of improved trajectory modeling and conflict probe performance due to increased entry of controller intent, using a fast-time simulation methodology that involves a CRA prototype and ERAM-like modeling environment. This study utilized a sound methodology utilizing several powerful tools and platforms including fast-time simulation airspace and ATC software, both internally developed and commercial off-the-shelf statistical and graphical platforms, and advanced multi-regression modeling to synthesize the results and estimate the net effects.

Overall, a performance improvement is observed in both trajectory modeling and conflict probe alerts with increasing levels of intent entry. Almost 45,000 flights over 240 hours are simulated and the output data is fit to a statistical model. The model fits the data closely, capturing between 97 and 100% of the variation in the data for different response variables, and indicates a strong non-linear effect from the parameter that reflects how completely controller intent is entered to the ground automation- the size of the effect is highest at lower levels of intent entry. The results indicate a potential improvement in trajectory modeling: a 61% decrease in the overall average horizontal error and a 42% decrease in the overall average vertical error when comparing scenarios that simulated the least amount of controller intent entry to scenarios with complete entry of full clearances. In addition, the number of trajectories generated that do not coincide with a flight plan amendment decrease by 42% overall between these scenarios, indicating that the trajectories generated are more stable and that less reconformance rebuilds are necessary with more complete entry of controller intent.

These improvements in trajectory prediction and more accurate trajectories lead to better performance of the conflict probe. Two metrics that are used to demonstrate conflict alert performance are the first quartile of predicted warning time and the count of alerts with duration greater than one minute. Inaccurate trajectories result in an increase in late notification of alerts, in other words, short warning time before a conflict. The authors consider the distribution of predicted warning time, or difference between predicted conflict start time and first time of notification for each conflict. The first quartile (25th percentile) of the predicted warning times of the alerts in seconds is a valuable metric to describe how the lower tail of the distribution is affected. From the statistical model, the first quartile of predicted warning time increases by 58 seconds overall when increasing full entry of 2-part clearances to the ground automation from 0 to 100%. The count of alerts with duration greater than one minute is important because with missing intent, some alerts will not be removed upon issuing an amendment. Alerts that are not deleted at the time an amendment is entered can be identified in this study by an alert duration greater than one minute, and represent problematic events. The count of these alerts decreases by an average of 80% over all experimental runs when increasing full entry of 2-part clearances from 0 through 100%.

To apply these results in estimating the impact of CRA, the percent of intent entry in current operations and future operations with CRA needs to be calculated. The percent of intent entry is related to the level of usage of CRA by ATC, and this can be estimated in a human-in-the-loop simulation or by surveying subject matter experts.

The full technical note documenting this study[12] presents detailed flight examples to demonstrate how the trajectory modeling and conflict probe alert performance is affected by the level of intent entry. Each example compares an instance from a reduced intent scenario, in which intent is not sent to the ground automation system, to the same time in the associated full intent scenario, with all clearances fully entered. The FAA FliteViz4D visualization tool[1] is used to explore the scenario data and produce graphics. The examples capture a wide array of benefit mechanisms from improved intent entry by demonstrating that in the simulated scenarios, when 2-leg maneuvers are not entered into the ground automation system, the modeled trajectories are incorrect and that this lack of correct intent information leads to false alerts, late or missed alerts, and unnecessary maneuvers. When the full clearances are entered, the trajectories are modeled correctly and these events are averted.

In summary, this study presents a comprehensive simulation of improved intent entry and evaluates the impact that Conflict Resolution Advisories could have on the performance of the ground automation with respect to trajectory modeling and conflict alert generation. The authors employed metrics that reflect important performance aspects of trajectory modeling and conflict probe alerting. The results indicate a significant impact and definite trend of performance improvement with increasing entry of full 2-part clearances.

Acknowledgments

The authors thank Karl Meyer and Dan Kirk of the MITRE Corporation's Center for Advanced Aviation System Development for their support of the fast-time simulation. Husni Idris and Vincent Kuo of Engility provided flight examples and analysis of benefit mechanisms. The authors also thank Al Schwartz, Jessica Young, Michael Konyak, and Doug Frye of ANG-C41 and Brian Schnitzer and Robert Oaks of General Dynamics Information Technology for reviewing and providing valuable feedback for the technical note and paper.

References

1. Crowell, A. T., Fabian, A. J., Nelson, N. A., *An Interactive 4D Visualization System for Air Traffic Concept Analysis*, 31st Digital Avionics Systems Conference, October 15-18, 2012.

2. Crowell, A. T., Fabian, A. J., Young, C. M., Musialek, B., Paglione, M. M., *Evaluation of Parameter Adjustments to the En Route Automation Modernization's Conflict Probe*, FAA Technical Note, DOT/FAA/TC-TN12/2, 2011.

3. Federal Aviation Administration, *TBO - Conflict Resolutions Advisories Benefits Plan, Version 0.1*, ATO-P, FAA Simulation and Analysis Group, William J. Hughes Technical Center, October 1, 2010.

4. Groupy, J., and Creighton, L., *Introduction to Design of Experiments with JMP® Examples, Third Edition*, Cary, NC, SAS Press Series, 2007.

5. Kirk, D.B., Arthur, W., Malik, Z., Meyer, K., and Steinbach, D., *Initial Trajectory Information Exchange Analysis* (Draft), The MITRE Corporation, McLean, VA, July 2012.

6. Kirk, D.B., Rozen, N.E., Waters, T., Arthur, W., and Meyer, K.A., *Problem Detection & Resolution Functional Performance Assessment – Plan and Interim Analysis Results, Revision 4,* MTR110414, The MITRE Corporation, McLean, VA, 2011.

7. Kuo, V. and Idris, H., *Safety Benefit Mechanisms for Improved Intent Using Initial Conflict Resolution Advisories*, FAA Task Order CA1513, Engility Corporation, Billerica, MA, June 2012.

8. Montgomery, D. C., *Design and Analysis of Experiments, Fourth Edition*, New York, NY, John Wiley & Sons, 1997.

9. Paglione, M. M., Oaks, R. D., *Implementation and Metrics for a Trajectory Prediction Validation Methodology*, American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation, and Control Conference, Hilton Head, S.C., August 2007.

10. Paglione, M. M., Oaks, R. D., and Bilimoria, K. D., *Methodology for Generating Conflict Scenarios by Time Shifting Recorded Traffic Data*, AIAA 2003-6724, American Institute of Aeronautics and Astronautics (AIAA) Aviation Technology, Integration, and Operations Conference, November 17-19, 2003.

11. Young, C. M., *Project Plan for CRA Fast-Time Simulation Study II Investigating Benefit from Improved Intent Entry*, FAA Memorandum, January 3, 2012.

12. Young, C. M., Fabian, A. J., Kuo, V., and Idris, H. R., *Trajectory Based Operations Conflict Resolution Advisories: Fast-Time Simulation Study Investigating Benefits from Improved Intent Entry*, FAA Technical Note, DOT/FAA/TC-TN12/50, November 2012.