

Determination of Lateral and Vertical Adherence to Route

Nicole Nelson* and Mike Paglione†

Federal Aviation Administration, Concept Analysis Group

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

Abstract

This paper presents an algorithm to evaluate the adherence of an aircraft to its known clearance. There are two independently operating algorithms involved in the determination of adherence, one evaluating the vertical dimension for adherence and another evaluating the lateral dimension. Both algorithms have recently taken on new functionality. This paper presents the algorithms, the recent modifications, and the study conducted to calibrate the parameters of the algorithms.

I. Introduction

Most air traffic service providers (ATSPs) across the globe continue to expect significant growth in air traffic demand in the future. If no action is taken, it is generally accepted that this growth will outpace the capacity limits of their aviation systems, resulting in greater congestion and inefficiency. In areas of the northeastern United States as well as Western Europe, these conditions may already have reached their capacity limits under peak demand. In unprecedented proportions, industry and ATSPs have responded by developing comprehensive plans requiring broad advances in ground-based and airborne automation. Most of these initiatives are well underway.

In the United States, the interagency Joint Development Planning Office (JPDO) foresees a traffic demand increase by 2025 up to three times the number of flights of today's traffic¹. The JDPO, as established in their charter under the "Vision-100" legislation (Public Law 108-176) signed by President G. W. Bush in December 2003, has mandated a next generation operational concept of the National Airspace System (NAS) for 2025¹. This next generation NAS envisions a trajectory-based separation management system that requires precise management of the aircraft's current and future position. The separation function of today, relying heavily on the cognitive skills of the air traffic controller to visualize aircraft trajectories on the radar display and issue resolutions via voice instructions to pilots, will be replaced by a distributed system of separation management components implementing performance-based separation standards. This future system will rely heavily on enhanced automation with conflict resolutions that are communicated digitally between air and ground and between aircraft.

A key automation component promoted in the JPDO's operational concept is the development of decision support tools (DSTs). These tools are envisioned to help mitigate many of the capacity and workload constraints of the system if effectively integrated with advanced automation solutions in the air and ground systems. These tools have many purposes and typically serve to reduce the cognitive workload faced by the current human decision makers operating the system. They include tools that serve to predict future conflicts between aircraft, both for ground based controllers or airborne pilots, allowing more strategic separation management of aircraft. Air traffic management DSTs include capabilities that forecast where and when traffic workload would stress the system. This allows air traffic supervisors to make more efficient adjustments to either avoid the condition or alter staff and/or

*Computer Scientist, FAA NextGen Concept Analysis Branch, ANG-C41, Nicole.Nelson@faa.gov, AIAA member

†Manager, FAA NextGen Concept Analysis Branch, ANG-C41, Mike.Paglione@faa.gov, AIAA senior member

airspace accordingly. Such tools also include air traffic metering tools to efficiently sequence aircraft into en route and arrival flows, maximizing the capacity of the system. A common thread in all these DSTs is the accurate and timely modeling of the aircraft's current state and anticipated future path. This modeling function is referred to as the trajectory predictor (TP) process.

A. Trajectory Predictor's Accuracy Impact to DST

The aircraft trajectory is the predicted 4-dimensional path from its current position to its planned destination. TP accuracy can be measured by post flight comparisons of predicted and observed aircraft trajectories. Since the predicted trajectory is the fundamental input that sustains the DST's capabilities and functions, the accuracy of the trajectory prediction has a direct impact on the DST's overall performance and usability. In order to attain the specified accuracy requirements of a DST, it is necessary to validate the TP. Ref. 2 presents a TP validation methodology that can drive the performance of a TP toward a targeted level. Ref. 3 defines system metrics used within this methodology and shows how these metrics can assess a TP's impact on a DST. For example, the temporal or longitudinal error associated with a TP's aircraft trajectory predictions will have a direct impact on the stability of a time-ordered schedule output from a metering DST. If a flight actually arrives significantly later than predicted, the DST's estimated time of arrival and associated order in the metering list will need to change. If the changes are frequent and sufficiently large, the utility of the generated schedule and the entire metering function will come into question.

As detailed in Ref. 3, the accuracy of the TP can be measured by post flight comparisons of predicted and observed aircraft trajectories. The TP requires many inputs to produce an accurate trajectory prediction such as wind and temperature forecasts, aircraft model characteristics, surveillance position reports, and flight path intent information⁴. Input factors of TPs have been the subject of many scientific studies. In Ref. 5, the National Aeronautics and Space Administration (NASA) ran aircraft field tests to verify the operational performance of its own TP. In a different study, researchers at the MITRE Corporation developed models to evaluate their DST's overall performance by utilizing accuracy statistics of their TP's performance⁶. In yet another effort, a collaborative group of European and American researchers illustrated that the impact of variations in these factors has significant effects on the output trajectory's accuracy⁷.

B. Flight Example of Missing Intent

Under present-day operations, the flight plan message is the typical means of coding both the aircraft operator's request and air traffic control's clearance of the aircraft's intended path. However, as the aircraft actually executes these maneuvers, unforeseen conditions such as the weather or the action of other aircraft, may impact the flight and require changes to the operation. These dynamic changes are often not processed the same by the current automation systems on the ground and on-board the aircraft. As a result, these systems are often not synchronized with respect to aircraft information.

A common example is the heading vector. To safely avoid other aircraft ahead, the current procedure is initiated verbally through direct radio communications between pilot and ground controller. Either to add delay or spatial distance to the aircraft's path, the air traffic controller instructs the aircraft pilot to deviate from the previously cleared flight plan to an alternate path. A specified heading is given for an indeterminate time or to capture a downstream position on the original flight plan. This information, although confirmed verbally between controller and pilot, is often not digitally transcribed for the automation on the ground. The result is aircraft predictions with missing lateral intent in the ground automation.

Heading vectors are not the only example of situations where ground automation lacks the clearances just issued to an aircraft. Flights may be verbally cleared to proceed direct to a downstream fix along its flight plan, presumably cutting time and distance off its overall route for improved efficiency and fuel savings. In the United

States, MITRE Corporation published a study in 2000 that reported that only about 30% of the lateral maneuvers within an en route facility were entered into the Air Traffic Management (ATM) automation⁸.

In other cases, the flight may be deviated to fly one or more hold maneuvers or parallel offset from the current route. This next example describes a flight entering a hold maneuver. An operational recording was made of a civilian airliner traveling into the United States' Washington Air Route Traffic Control Center (ARTCC), referred to as ZDC. It originated from Hartsfield-Jackson Atlanta International Airport (ATL) in Atlanta Georgia with the destination of Reagan National Airport (DCA) in Arlington Virginia.

Figure 1 displays the three components of the flight: the expanded flight plan route is represented by a solid line, the predicted trajectory is displayed as a three dimensional wire frame, and the actual aircraft's path, captured from the surveillance radar track, is displayed as a series of points. The focus of this example is the ground automation's trajectory built at 74018 seconds (20:33:38 UTC). Of particular interest is the complete hold maneuver performed later in the flight beginning roughly at 20:40 UTC represented graphically as a series of circular arcs produced by the plotted track points. Clearly, the trajectory does not reflect this event, which is suspected to be a result of a verbal air traffic control clearance not entered into the automation system.

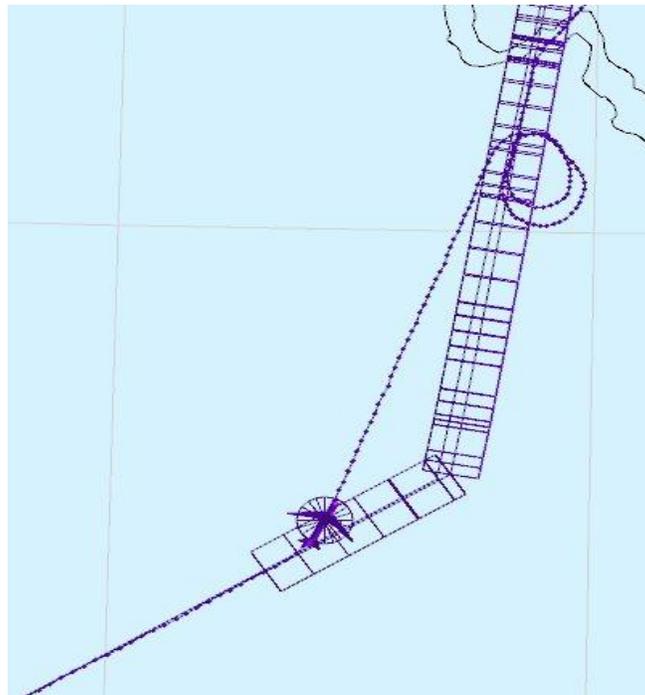


Figure 1. Close-up View of Actual Versus Trajectory

Table 1 lists the calculated trajectory metrics, comparing the trajectory to the recorded track data, for the 74018 second trajectory. A sample was taken at 74040 seconds (20:34:00) with a look-ahead time every five minutes up to 15 minutes in the future. At the first measurement time at look-ahead time of zero, the horizontal error (i.e. straight-line unsigned error) was less than a tenth of a mile and the vertical error was 55 feet. However, as the look-ahead time progresses and approaches the turn the horizontal error increases significantly. Due to the missed maneuver, the error reaches up to nearly 38 nautical miles (nmi) horizontally. The clearly visible cross-track error (i.e. side-to-side lateral error) is approximately 6 nmi, yet the bulk of the error is found in the along-track error (i.e. longitudinal or along the route error), up to 36 nautical miles.

Table 1. Sample Flight's Trajectory Metrics

Measurement Time (HH:MM:SS)	Look-Ahead Time (Seconds)	Horizontal Error (nmi)	Cross-track Error (nmi)	Along-track Error (nmi)	Vertical Error (Feet)
20:34:00	0	0.8	-0.07	-0.05	55
20:39:00	300	1.4	-1.3	-0.2	-916
20:44:00	600	26.5	6.2	-24.9	1768
20:49:00	900	37.9	5.7	-36.2	7657

The intent error is not limited to the horizontal dimension. Figure 2 illustrates the time versus altitude profile of the aircraft. The trajectory is built with a clearance posted to descend to 10,000 feet, posted at approximately 20:33 UTC. The aircraft performs the descent and is instructed into a holding maneuver upon reaching 10,000 feet. The trajectory predicts the continuation of descent into the arrival airspace of DCA based on an altitude restriction. However, since the aircraft is held at 10,000 feet the TP has difficulty in accurately estimating when the aircraft will continue the descent. Like the horizontal dimension where the TP lacked the knowledge of the holding maneuver, the error based on time increases rapidly. In this example, Table 1 shows the vertical error quickly grows to nearly 8000 feet by 20:49:00 UTC.

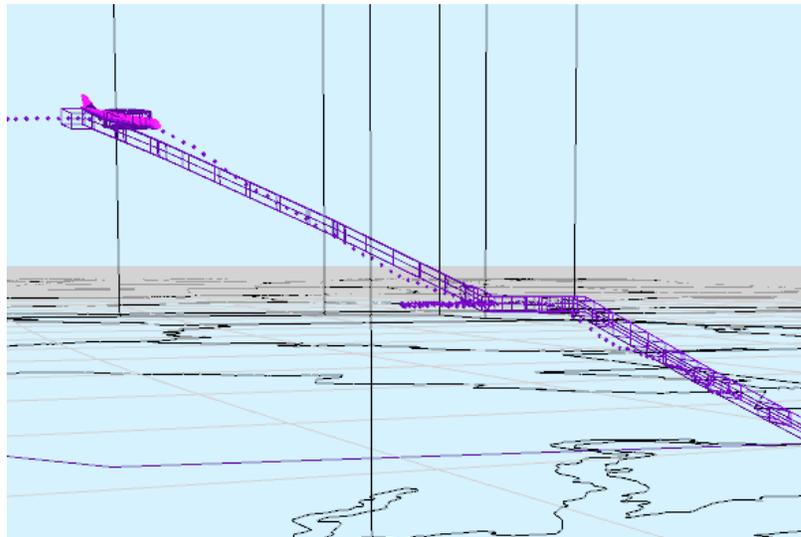


Figure 2. Sample Flight Altitude Plot

Clearly, the performance of the TP and later DST functionality will be quite different if aircraft maneuvers in the form of heading vectors, holds, or changes in the lateral or vertical path of an aircraft are not provided to the ground based TP. The example demonstrated how these events can cause large errors in the predictions of the trajectory. In the next section, methods will be described that will identify significant lateral and vertical events, allowing the analyst to separate these situations from nominal conditions and therefore better estimate the overall performance of the TP.

II. Adherence Detection Algorithms

The concept of adherence includes both the aircraft's lateral proximity to its route and its vertical proximity to posted altitude clearances. Although the algorithms together define adherence, they work independently. For this reason, the algorithms and modifications are described separately in the following subsections.

A. Lateral Adherence Detection Algorithm

Figure 3 shows a flowchart representing the functional process of the lateral adherence algorithm. The flowchart is comprised of rectangular objects representing different states of the algorithm, rounded rectangular objects representing the decision made by the algorithm, and unidirectional arrows depicting the flow from state to state. The lateral adherence algorithm depends on multiple parameters that are discussed further in the Lateral Adherence Parameter Calibration section.

The algorithm utilizes the calculated lateral distance ($latDist$) from the current track point to the route. The lateral distance is first checked against the inner threshold to determine if the aircraft is currently close to its route. If the aircraft is close, within the inner threshold distance, the aircraft is determined to be in adherence without considering any other factors.

When the aircraft is not considered close to its route, the algorithm detects if the aircraft has reached and exceeded the end of its route by the inner threshold distance. As long as the aircraft is within the inner threshold distance of the end of the route, the aircraft is still considered in adherence.

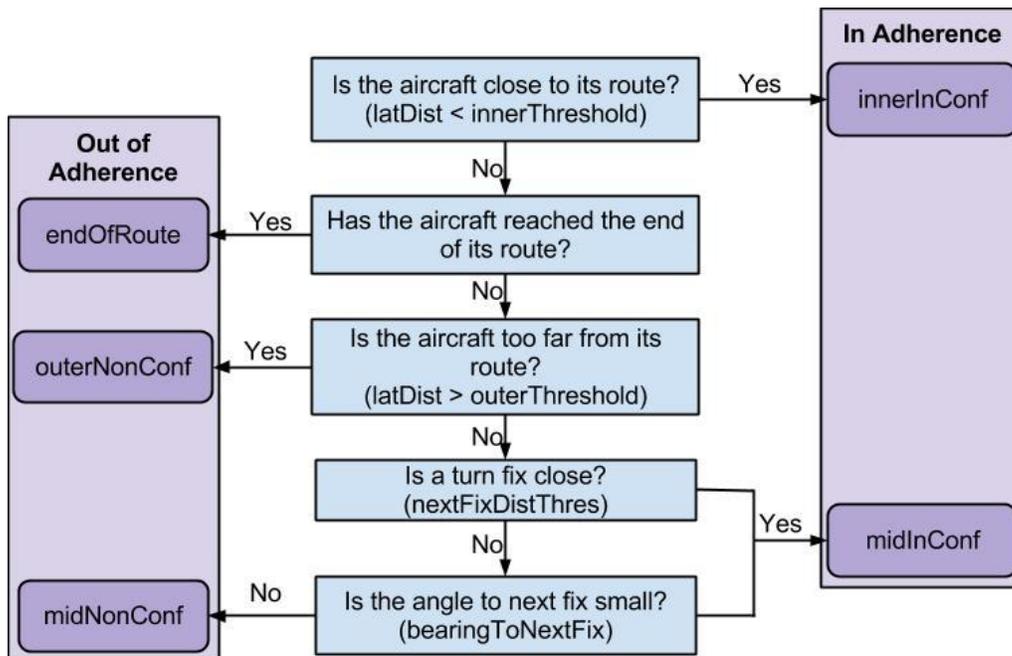


Figure 3. Lateral Adherence Algorithm Flowchart

The outer lateral threshold is applied before conducting any further tests in the algorithm since an aircraft with a lateral distance greater than the outer lateral threshold is considered to be out of adherence under all circumstances. If the lateral distance is not greater than the outer lateral threshold then it must be true that the aircraft has a lateral distance that lies between the inner lateral threshold and the outer lateral threshold.

When an aircraft's lateral distance is between the two thresholds, it must exhibit the intent of returning to its route in order to be considered in adherence. Next, the aircraft's position is used to determine if it is near a turn fix. A turn fix is defined as a fix with an angle, consisting of rays pointing toward the previous and next fix and vertex equal to the current fix, with measure greater than a threshold indicating a turn in the route. Specifically, if a fix indicating a turn exists is within a threshold distance to the current track position, the aircraft is considered in adherence.

When the aircraft is not near a turn fix, the track bearing angle is calculated between the aircraft and the next fix on the route. If the angle is within the bearing to next fix threshold it is concluded that the aircraft is heading back to its route. In this case it is considered to be in adherence; otherwise, it is out of adherence.

For lateral adherence determination, the change in algorithm does not pertain to the flow of the algorithm but rather to a portion of the algorithm. The algorithm relies on the angle to next fix under certain conditions as previously discussed.

Figure 5 depicts the original algorithm for determining angle to next fix. The algorithm first finds the route segment which the current track position is closest to. When calculating angle to next fix it simply uses the end node of the closest segment. In Figure 4, the current track point is numbered 5. The closest segment is the line connecting Node1 and Node 2, so, the angle to next fix is calculated on Node 2.

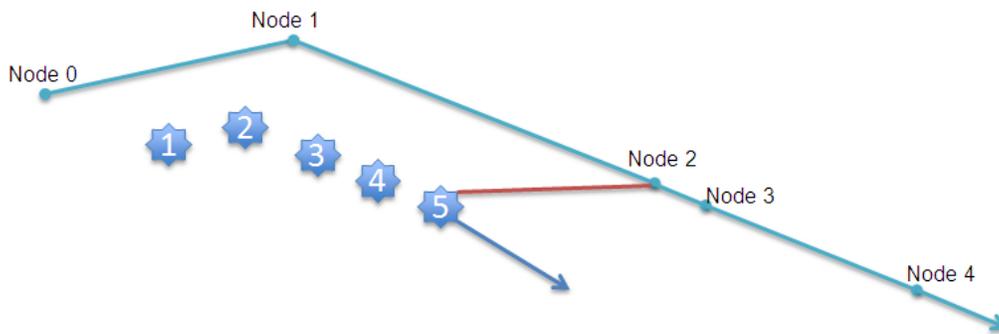


Figure 4. Original Next Fix Angle Calculation

By the definition, as the aircraft approaches the next fix, this bearing angle that is formed will get larger. As a result, the authors determined that the logic for calculating angle to next fix needed to be reworked. The goal was to develop an algorithm for choosing the next fix which was more robust and did not grow in magnitude as the flight progressed in time and moved closer to the targeted next fix.

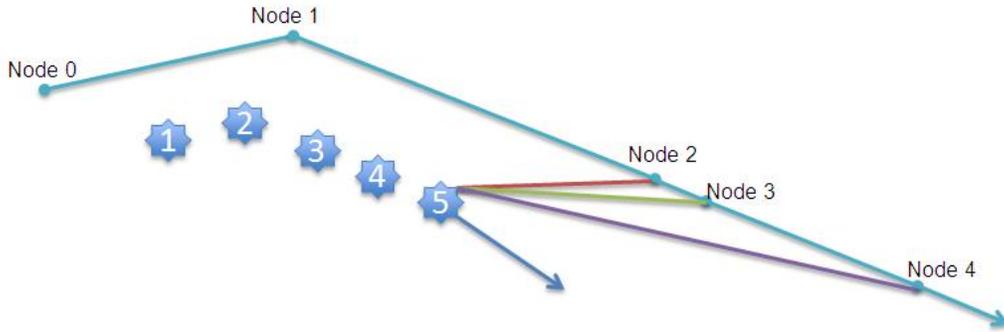


Figure 5. Adjusted Next Fix Angle Calculation

Figure 5 displays the modified algorithm used for determining which fix should be identified as the next fix for angle calculation. The algorithm still begins by determining which segment is the closest segment, as the original algorithm did. Then, rather than just applying the second node of the segment, the algorithm includes a proximity check between the current track position and the fix.

The modified algorithm calculates the distance to multiple fixes for each track point in order to find a fix that is further than the distance threshold away. In Figure 5, the aircraft is again currently at the position identified by 5. Now, however the algorithm checks the distance between the current position and Node 2. The distance is determined to be too small so the algorithm checks the position against Node 3. Again the distance is too small. When the algorithm checks the proximity to Node 4, the distance is greater than the threshold. The next fix angle is calculated based on Node 4, in this case, which is the first fix greater than the distance threshold.

This modification allows the resulting adherence to more closely relate to the adherence determined through both numerical and visual analysis. The goal of the algorithm is to perform in the same manner as human detection, resulting in the same in or out of adherence determination.

B. Vertical Adherence Detection Algorithm

Figure 6 is a flowchart representing the functional process of the vertical adherence algorithm. Each rectangular object represents a process performed when determining adherence in the vertical dimension. Again rectangular objects represent states, unidirectional arrows indicate the flow, and rounded rectangular objects represent endpoints in the adherence determination. The actual flow of the algorithm will be described in the following paragraphs.

The major change in the algorithm for determining vertical adherence is the ability to make use of previously cleared altitudes. The new functionality allows an aircraft to be in adherence with a previously cleared altitude when it would be out of adherence with its currently cleared altitude. The altitude applied by the algorithm is referred to as the probed altitude, whether it is the currently cleared altitude or a previously cleared altitude.

Including the new functionality of probing on previously cleared altitudes also required some additional logic in the algorithm for determining which altitude to apply and when an altitude can no longer be used. A new set of rules was implemented to maintain the collection of clearances and for probed altitude determination.

The first rule determines that a point is out of adherence when it is diverging from its currently cleared altitude and was level at its previous track point. The next two rules pertain to maintaining the list of previously cleared altitudes. First, the currently cleared altitude is added to the list if it is not already the last element in the list of clearances. Second, when an aircraft is in adherence, any clearance that came before the probed altitude can no

longer be considered for adherence determination. So, when the aircraft is in adherence with a given clearance everything before it in the collection of clearances is removed.

Lastly, a determination needs to be made regarding how an altitude should be selected for use in probing. Beginning with the first clearance in the collection, adherence is checked against each clearance in the list until one is found that is in adherence or until each element in the collection has been checked. In the case that a clearance is found that will result in the flight being in adherence, another check is made to see if the track is only in conformance based on the vertical conformance threshold. Consider the case where the vertical conformance threshold is 300 feet and the probed aircraft is ascending. When determining the probed altitude, a clearance is found where the aircraft is above and within 300 feet of the clearance and the aircraft is still below but ascending towards the next clearance. In this case the probed altitude will be the second clearance which the aircraft is currently below but ascending toward. Similarly, this logic is also applied to descending aircraft.

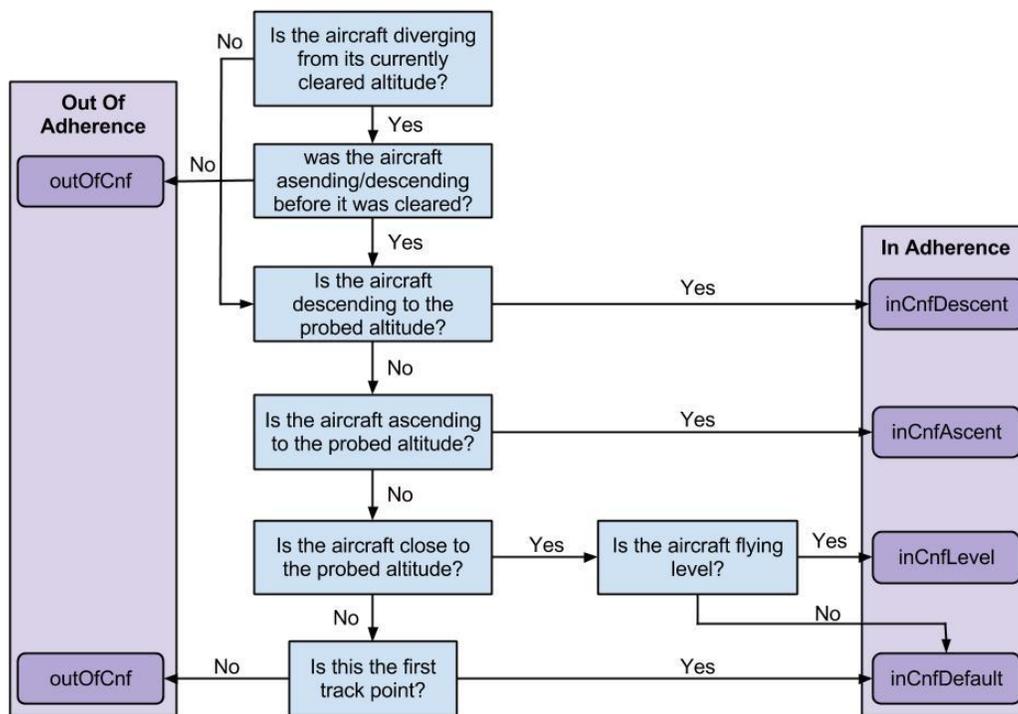


Figure 6. Vertical Adherence Algorithm Flowchart

The first step in determining vertical adherence is determining the altitude to probe at. The flowchart depicts this by first checking the divergence rule; if the aircraft begins diverging from its current clearance after the clearance is posted it is automatically determined to be “outOfCnf”. If the aircraft is not diverging, the probed altitude is chosen from the previous clearance collection. In the determination of vertical adherence, multiple parameters can change the result. These parameters are discussed in detail in the Vertical Adherence Parameter Optimization section later in this paper.

The algorithm continues with the probed altitude by first determining if the aircraft is ascending, descending or level. This determination is dependent on parameters. If the aircraft is ascending to a probed altitude above its current altitude it is “inCnfAscent”. The descent logic is just the opposite, to be “inCnfDescent” the aircraft must be descending to a probed altitude that is less than its current altitude.

If the aircraft is neither “inCnfAscent” nor “inCnfDescent”, the algorithm performs a comparison between the current altitude and the probed altitude. If the aircraft is within a 300 feet of its probed altitude it can either be considered “inCnfLevel” or “inCnfDefault”. The aircraft is “inCnfLevel” when the provided parameters determine the aircraft is flying level within 300 feet of its probed altitude or “inCnfDefault” when it is within 300 feet of its probed altitude but is not flying level. When the probing is performed for the first track point, the aircraft can be considered “inCnfDefault” if it doesn’t fall into one of the previously mentioned categories. If at the current track time the aircraft does not fall into any of the cases discussed above, the aircraft is considered to be “outOfCnf”.

III. Adherence Parameter Calibration

Fine tuning an algorithm’s parameters is an important step in algorithm development. Although the majority of the parameters used in testing existed prior to the recent changes, it is still important to find their optimum settings based on the updated algorithms. The details of the Design of Experiment (DOE) conducted for the two algorithms are presented in the next section. Separate DOEs are performed for the lateral adherence algorithm and the vertical adherence algorithm because, as previously mentioned, the two algorithms function independent of one another.

A. Design of Experiment Metrics

Although the vertical and lateral adherence algorithms perform differently and independently, the overall concept of their design is the same. The vertical algorithm determines adherence based on altitude clearances, while the lateral algorithm determines adherence based on proximity to the cleared route from the flight plan. Based on the similarity of the algorithms, the DOE follows the same model for each. Another reason two experiments were conducted separately was to reduce the size of the experiment considering the large number of factors that are required to effectively study both algorithms in one experiment.

Table 2. Metric Determination Events

		Signal Adherence	
		Out	In
Response Adherence	Out	Valid Call (VC)	False Call (FC)
	In	Missed Call (MC)	No Call (NC)

The same basic metrics were calculated and applied in determining the optimal parameter settings. In both algorithms, the actual determination for in or out of adherence, the response, is compared to the signal generated manually through observation of the geometry. This comparison results in four independent events, displayed in Table 2. The four cases are: Valid Call (VC), False Call (FC), Missed Call (MC), No Call (NC). A valid call occurs when the response and the signal both determine out of adherence. A false call occurs when the response determines out of adherence, but the signal indicates in adherence. A missed call occurs when the response determines in adherence, but the signal indicates out of adherence. A no call occurs when both the signal and response indicate in adherence.

Each aircraft may have a large number of each event type occurring through the duration of flight as adherence is calculated for each surveillance radar track point, with track points interpolated to every 10 seconds. For this reason, the different events were not used directly, but rather to define the metrics used for optimization of algorithm parameters. The metrics used to determine optimal parameter settings were missed call probability (P(MC)), false call probability (P(FC)), signal's nearest response time (SRT), and response's nearest signal time (RST).

The first two metrics, missed call probability (P(MC)) and false call probability (P(FC)), are determined based on the number of occurrences of the defined events. P(MC) is represented in Eq. (1) and is defined as the ratio of missed calls to signal out of adherence events, where actual events are the total number of missed calls and valid calls. Eq. (2) provides the definition of P(FC): the ratio of false calls to signal in adherence events, where these non-signal events are the total number of false calls and no calls.

$$P(MC) = \frac{\text{number of missed calls}}{\text{total number of signal out of adherence events}} = \frac{\text{number of MCs}}{\text{number of MCs} + \text{number of VCs}} \quad (1)$$

$$P(FC) = \frac{\text{number of false calls}}{\text{total number of signal in adherence events}} = \frac{\text{number of FCs}}{\text{number of FCs} + \text{number of NCs}} \quad (2)$$

The remaining two metrics, Signal's nearest Response Time (SRT) and Response's nearest Signal Time (RST), evaluate the degree of timing error. These metrics require slightly more explanation. Put simply, these represent the average absolute values of the amount of time for a call to be made which would create a valid call, referred to as a corresponding call. The two metrics are similar in design, varying based on the triggering event type and search space. The triggering event is the event found before calculating the respective metric.

For SRT a missed call is the triggering event which then uses the response as the search space to find the nearest corresponding call before or after the missed call time. Referencing Figure 7, a missed call is detected at time 20. In searching, a corresponding response call is detected at time 60 which results in a 40 second SRT for this track point. A false call is the triggering event for RST, which then uses the signal as the search space to find the nearest corresponding call that occurs before or after the false call time. Again referencing Figure 7, a false call is detected at time 60. A corresponding signal call is found 40 seconds earlier at time 20, resulting in a 40 second RST for this track point.

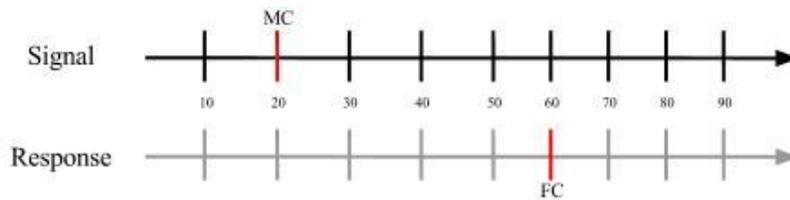


Figure 7. SRT and RST Metrics

The search space is limited to within 60 seconds of call time. When a corresponding call is detected before the call time a negative time is stored, and a positive time results for detection after the call time. In the case that a corresponding call is found the same amount of time before and after, the positive time is stored. If no corresponding call is found within 60 seconds, the time is stored as 60. For use in this study, the absolute value of RST and SRT were utilized.

Logically, P(FC) and RST are related because they are both driven by the false call events with P(FC) providing the ratio of false events to all alerting opportunities (i.e. total false calls and correct no-calls), and RST response metric being the mean time that a signal was near these false call events. Similarly, the P(MC) and SRT

are related. $P(\text{MC})$ is the ratio of missed calls to the total available signals available to call (i.e. total missed calls and valid calls), and SRT response metric is the mean time that response is near a missed call. Thus, $P(\text{MC})$ and SRT are driven by the missed call events.

The optimal solution will minimize all four metrics. Ideally the algorithm would produce only no calls and valid calls. The best solution should minimize false calls and missed calls. It is also ideal for the SRT and RST to be small time values. These values indicate how close in time the missed call or false call came to the correct prediction. A zero SRT or RST maps to a valid call.

B. Truth Reference Data Set

In order to produce the metrics discussed in Section III.A there needs to be a signal to compare the response to. In order to collect data for the signal, a scenario of traffic data needed to be identified for analysis. The selected scenario comprises a total of 100 flights which were semi-randomly selected from a 2,234 flight scenario in the Chicago Air Route Traffic Control Center. The selection was based on certain criteria: 40% of the aircraft had to have at least one point out of adherence vertically and 40% of the aircraft had to have at least one point out of adherence laterally. The remaining 20% were selected at random resulting in flights that have zero or more points out of adherence.

Due to the complexity of the lateral adherence algorithm the data set was reduced to a smaller subset of 50 flights selected based on the effect that the algorithm would have. If aircraft are far from their route or follow their route for the entire flight the algorithm would not have the opportunity to change the adherence determination, so these flights were removed from the data set.

Each track position then was evaluated to determine the state (in or out) for both lateral and vertical adherence. This determination acts as the truth data, which serves as the signal value used to make evaluations based on event detection and metric calculation of the response generated in each run indicated by the model.

C. Design of Experiments

An experiment is defined in Ref. 9 as “a test or series of tests in which purposeful changes are made to input variables of a process or system so that we may observe and identify the reasons for changes in the output response.” Experiments are utilized for testing and determination by researchers and scientists from practically all disciplines by researchers and scientists.

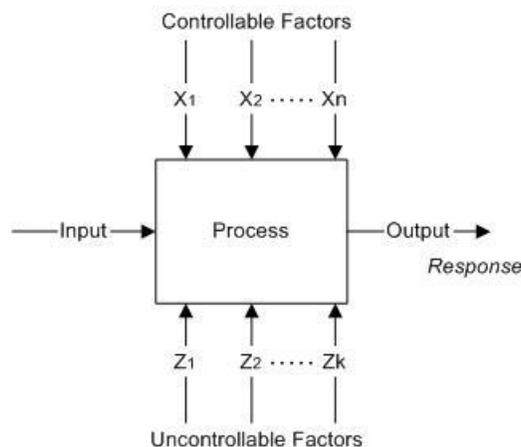


Figure 8. General Model of a Process (adapted from Ref. 9)

Figure 8, extracted from Ref. 9, represents the functional flow of an experiment. An experiment is made up of an input, controllable factors, uncontrollable factors, an output, and a process to get from the input to the output. The experimental input consists of the actual input and the controllable factors. These controllable factors are the factors that the model manipulates to generate a response. In addition, uncontrollable factors exist which are not easily manipulated. However, uncontrollable factors can be removed through experimental design techniques such as blocking and randomization. The output or response is the data generated by the process using the determined settings for the controllable factors. Results related to responses are then studied and analyzed.

The purpose and goal of the experiment performed for this study is to determine the parameter settings for the Lateral and Vertical Adherence Algorithms which would best calibrate the algorithms. The solution for the parameter settings will attempt to minimize the four metrics defined in Section III.A. There are seven controllable factors for the lateral adherence algorithm, seen in Table 3. The vertical adherence algorithm utilizes four controllable factors defined in Table 7.

In the implementation of these experiments, a model was defined with the objective of accurately testing and evaluating the factors. The general model of this experiment is expressed in Eq. (3).

$$\gamma = \mu + \tau + \varepsilon \quad (3)$$

*where γ is the response, τ is the effect,
 μ is the constant, and ε is the error term*

The design specific to these experiments includes the two-way interactions of the factors defined for the algorithm being tested. These factors and interactions contribute to the effect portion of the design, τ . In general, τ is defined by Eq. (4).

$$\tau = \text{main effects} + \text{interaction effects} + \text{nonlinear main effects} \quad (4)$$

This equation is expanded to detail the effects included in the experiment. In an experiment with 3 factors, A , B , and C , $\tau = A + B + C + AB + AC + BC + A^2 + B^2 + C^2$. The example displays the main effects as single letters, two-way interactions as pairs, and polynomial effects as squares. The expanded definition of τ for each experiment is defined in its respective subsection below.

1. Lateral Adherence Parameter Calibration

The Lateral Adherence Algorithm has seven parameters as illustrated in Figure 3. These parameters and their definitions can be found in Table 3. There are two types of parameters: continuous and fixed. Continuous parameters are bound by a range for the purpose of this analysis. Fixed parameters are presented with a specific set of values. For the Lateral Adherence Algorithm, there are only two fixed parameters, *changeStateToLatOut* and *changeStateToLatIn*. The remaining five parameters are continuous.

Table 3. Lateral Adherence Algorithm Parameters

Parameter Name	Parameter Description	Related Model Values
<i>innerLatThres</i>	These properties are decimal values, in nmi, used to determine the range category of the lateral distance between the route and track. These properties determine which other properties, if any, need to be evaluated when determining adherence.	0.25 – 1.5 nmi
<i>outerLatThres</i>		1.5 – 3.0 nmi
<i>changeStateToLatOut</i>	These input properties specify the number of consecutive points that need to be flagged as laterally in or out of adherence, respectively, before the algorithm declares the point in or out, respectively.	1, 2, 3 points
<i>changeStateToLatIn</i>		1, 2, 3 points
<i>bearingToNextFix</i>	This input property specifies an angle, in degrees, between an aircraft and its next fix which is used to indicate when an aircraft is returning to its route.	15.0 – 45.0°
<i>nextFixDistThres</i>	This input property is a decimal value representing the range of nmi in which to look for the next fix, starting from the current track position point.	3.0 – 15.0 nmi
<i>longDistThres</i>	This input property is a decimal value representing a range of nmi, from the current track position point, within which a turn fix may be included in the adherence determination.	3.0 – 20.0 nmi

A full factorial experiment exercises all combinations of factor level. For the Lateral Adherence Algorithm this results in 288 (2 x 2 x 3 x 3 x 2 x 2 x 2) experimental runs plus an additional five runs to test for non-linearity in selected continuous factors. The four response variables were defined in Section III.0. The following equation provides an overall model for the Lateral Adherence Algorithm.

The treatment effect in the above model represents the full factorial of the seven factors under study. The full factorial yielding 288 runs contains all seven factors under study and all their interactions. For a seven factor experiment with two factors at 3-levels, the equation has 288 terms plus a term for the non-linear factor and is simply too large to publish here. For simplification, Eq. (5) contains the seven factors with their main and two-way interaction terms and a term for a selected non-linear factor. For a full explanation of this model and the assumptions involved see Ref. 9.

$$\tau = A + B + C + D + E + F + G + AB + AC + AD + AE + AF + AG + BC + BD + BE + BF + BG + CD + CE + CF + CG + DE + DF + DG + EF + EG + FG + C' + AC' + BC' + DC' + EC' + FC' + GC' + D' + AD' + BD' + CD' + ED' + FD' + GD' + B^2 \quad (5)$$

where *A* is the first factor, *B* is the second factor up to *G* which is the last factor, *C'* and *D'* represent the third setting for the third and fourth factor, and lastly the polynomial model for factor *B*

The fitted model is summarized graphically in Figure 9 for two of the four response variables. The leverage plot illustrates the actual and modeled values for a given response variable. 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 figure and the coincident modeled version on the x-axis would fall perfectly on a diagonal line. The term “Rsq” in the plot 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. A value of 1 indicates that 100 percent of the variation is captured, and a value of 0 that indicates none of the variation is captured.

For this experiment, all four response metrics had Rsq values near 1, indicating that close to all of the variation is being captured by the model. In Figure 9, the SRT response variable has an Rsq value of 0.99 and P(MC) has a value of 1.00.

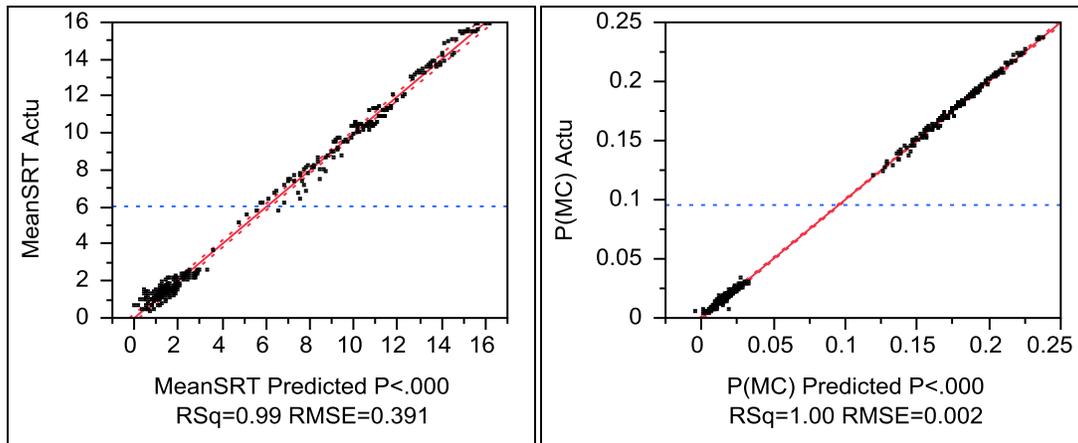


Figure 9. Lateral Adherence Leverage Plot for SRT and P(MC) Response Variables

Table 4 and Table 5 display the effect tests for the various factor level combinations of the experiment. They only provide results for one of the output variables, P(MC), but the other three look very similar. The full set of results for all four output variables is listed in Table 11 in the Appendix. For Table 4 and Table 5, 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 and subtracting the mean. The column labeled “F Ratio” is the test statistic produced by taking the model mean square and then dividing it by the error mean square. The column labeled “Prob > F” is the p-value, which 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 4. Lateral Adherence Model Main Effects for P(MC)

Source	DF	Sum of Squares	F Ratio	Prob > F
innerLatThres	1	0.0024	392.5295	0.0000*
outerLatThres	1	1.8955	315076.4524	0.0000*
changeStateToLatOut	2	0.0115	952.9227	0.0000*
changeStateToLatIn	2	0.0144	1197.6012	0.0000*
bearingToNextFix	1	0.0348	5785.8641	0.0000*
nextFixDistThres	1	0.0007	110.9064	0.0000*
longDistThres	1	0.0074	1234.1110	0.0000*

In Table 4 the main effects of all seven factors have p-values less than 0.05, indicating that all seven factors are statistically significant for P(MC). In Table 5, the analysis continues for all of the interaction effects and for the single polynomial effect for the *outerLatThres* factor. When considering the interactions, there some factor combinations that are significant for P(MC) and others that are not. Also, the polynomial *outerLatThres* factor effect (last row in Table 5) is significant as well, indicating that this factor has a non-linear impact on the P(MC) response variable.

Table 5. Lateral Adherence Model Two Way Interactions and Polynomial Effects for P(MC)

Source	DF	Sum of Squares	F Ratio	Prob > F
innerLatThres*outerLatThres	1	0.0000	0.0846	0.7714
innerLatThres*changeStateToLatOut	2	0.0001	6.5266	0.0017*
innerLatThres*changeStateToLatIn	2	0.0000	1.4026	0.2479
innerLatThres*bearingToNextFix	1	0.0006	92.3139	0.0000*
innerLatThres*nextFixDistThres	1	0.0000	0.2930	0.5888
innerLatThres*longDistThres	1	0.0002	35.1739	0.0000*
outerLatThres*changeStateToLatOut	2	0.0014	119.0223	0.0000*
outerLatThres*changeStateToLatIn	2	0.0071	586.0368	0.0000*
outerLatThres*bearingToNextFix	1	0.0223	3702.0866	0.0000*
outerLatThres*nextFixDistThres	1	0.0005	80.8495	0.0000*
outerLatThres*longDistThres	1	0.0038	630.7618	0.0000*
changeStateToLatOut*changeStateToLatIn	4	0.0001	3.9963	0.0037*
changeStateToLatOut*bearingToNextFix	2	0.0005	45.0019	0.0000*
changeStateToLatOut*nextFixDistThres	2	0.0000	0.1124	0.8937
changeStateToLatOut*longDistThres	2	0.0000	0.5356	0.5860
changeStateToLatIn*bearingToNextFix	2	0.0000	0.1449	0.8652
changeStateToLatIn*nextFixDistThres	2	0.0000	0.0266	0.9737
changeStateToLatIn*longDistThres	2	0.0000	0.2210	0.8018
bearingToNextFix*nextFixDistThres	1	0.0000	7.6557	0.0061*
bearingToNextFix*longDistThres	1	0.0001	14.7565	0.0002*
nextFixDistThres*longDistThres	1	0.0007	113.5757	0.0000*
outerLatThres*outerLatThres	1	0.0502	8351.6645	0.0000*

Figure 10 presents the predictor profile plot of the model results for the Lateral Adherence Algorithm. The JMP® commercial software tool generates an interactive model calculator called the predictor profiler. The predictor profiler illustrates the effects of the various factors of the model and is used during analysis to explore the relationships and settings of the various factors and responses.

In general, the profile confirms that the *outerLatThres* factor is the factor with the biggest impact on the predictions and that it is non-linear. The predictor profiler in Figure 10 illustrates that these response variables display similar patterns in terms of their proportional change to the various factor levels. For example and most notably is *outerLatThres* factor, where the P(MC) non-linearly increases as the factor level increases. From a factor level of 1.5 to a level of about 2.3 nmi, the increase in the P(MC) is shallow, and then it increases steeply until peaking at the maximum tested value of 3.5 nmi. A similar pattern is seen for the *outerLatThres* factor and SRT. For P(MC), logically as the threshold increases more signal is considered in adherence and is thus more likely to be categorized as missed. For SRT, as the threshold increases it is more likely that there will be a delay in detecting the signal, thus increasing the time to provide a response to detect the signal. The fact that both are non-linear indicates that below about 2.3 nmi, the *outerLatThres* factor is less responsive in detecting signal and other mechanisms of the algorithm are detecting signal. Above 2.3 nmi, the *outerLatThres* takes on a major role and is very responsive.

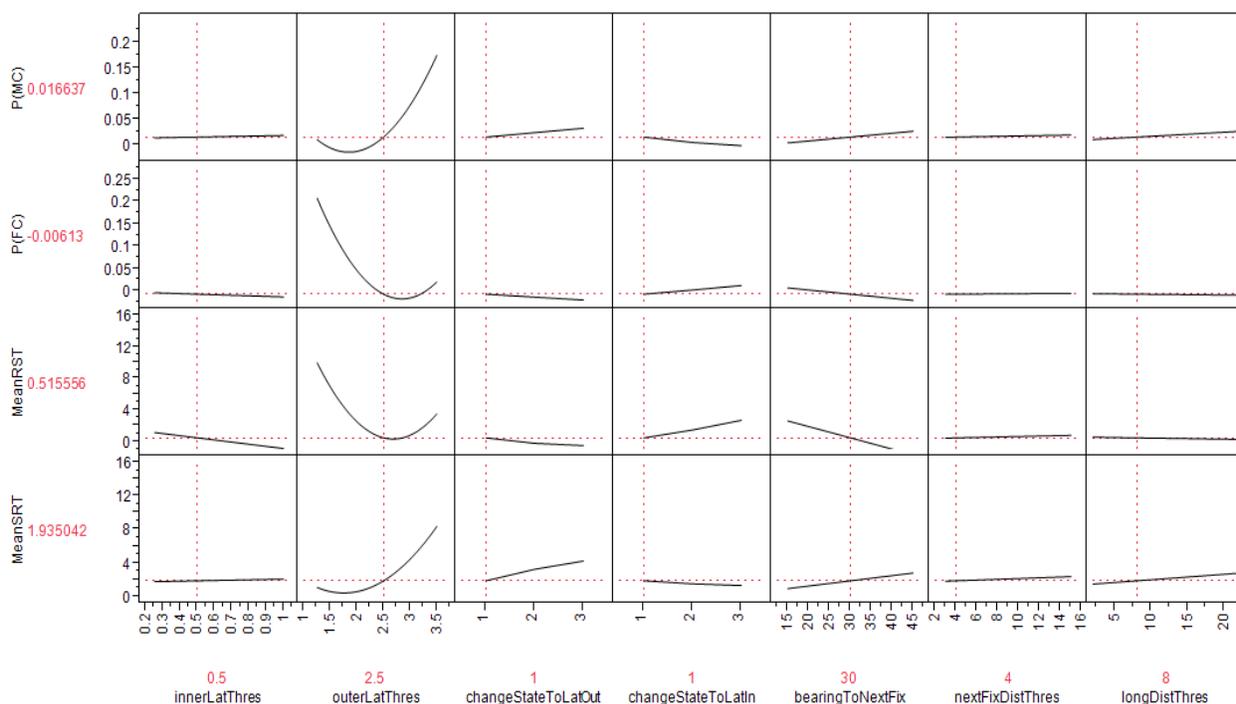


Figure 10. Lateral Adherence Model Results

The slope of the functions in the predictor profiler in Figure 10 indicates the magnitude of the effect of that factor over the range of its factor levels of the experiment. As expected the two fixed factors, *changeStateToLatOut* and *changeStateToLatIn*, have opposite and opposing effects on the detection performance. Logically this makes perfect sense because both are delaying or accelerating the detection process but from opposite points of view. The *changeStateToLatOut* factor requires the parameter number of continuous detections to change state from in to out of adherence, while the *changeStateToLatIn* factor requires the parameter-defined number of continuous non-detections to go back into adherence. As expected, increasing the “out” threshold increases P(MC) and SRT, and

the opposite is found for the *changeStateToLatIn* factor. Similarly the opposite relationships are found for P(FC) and RST. Interestingly, the largest noticeable effects (slopes in the predictor profiler in Figure 10) for these factors are for the *changeStateToLatOut* factor on the SRT output variable and the *changeStateToLatIn* factor on the RST output variable.

The remaining factors, *innerLatThres*, *bearingToNextFix*, *nextFixDistThres*, and *longDistThres*, all have impacts as expected. The largest of the four is *bearingToNextFix*. As *bearingToNextFix* increases P(MC) and SRT rise which indicates that as the angle is enlarged more aircraft track positions are evaluated to be in adherence. This increases the chances of not detecting and delaying detection of a signal of out of adherence. The opposite is illustrated for P(FC) and RST.

The factorial experiment performed, as defined by the model in Eq. (5), can also detect and model interactions between factors. There are many interactions between these factors. Most notably are the interactions listed in Table 5 between the *outerLatThres* factor and all of the other factors. This is expected to some degree given the interrelationships of the factors as defined earlier in the flowchart in Figure 3. What would have been difficult to determine without the experiment and resulting model is the magnitude of these impacts.

The values of the factors listed on the bottom of Figure 10 are not the optimal settings but are near optimal. They represent a combination of operational settings and optimal settings from the model with predicted values of good overall performance. To validate the model further, the parameter settings were actually run on the same test scenario and the results were compared to the model predictions. The model results are listed in Table 6. The resulting residuals are negligible. The only minor issue is the predicted value of P(FC). As illustrated graphically in the predictor profiler in Figure 10, the model dips slightly below zero which is unrealistic, since both P(FC) and P(MC) are probability metrics with a range of 0 to 1 by definition. However, the fact that it is only slightly negative does not negate the overall performance of the model. The authors conclude only slightly negative, since the “Rsq” metrics for all four response variables was calculated to between 1.00 and 0.99, indicating that most of the residuals must indeed be small.

Table 6. Model Results for Near Optimal Settings of Lateral Adherence Algorithm

	P(MC)	P(FC)	MeanRST (sec)	MeanSRT (sec)
Predicted	0.016637	-0.006135	0.515556	1.935042
Actual	0.016027	0.0033	1.05	1.517424
Residual	-0.00061	0.009435	0.534444	-0.417618

To further confirm the model’s validity, residuals for all the 293 experimental runs were calculated and examined in detail. All four response variables produced distributions that were approximately normally distributed with means near zero. For example, the model’s residuals of the SRT response variable is illustrated by histogram in Figure 11 and normal probability plot in Figure 12. Both provide a graphical diagnostic for testing whether a data sample matches a normal distribution. In this case, the residuals or random error that the model captures is approximately normally distributed, a major assumption in the original additive model defined in Eq. (3) presented earlier.

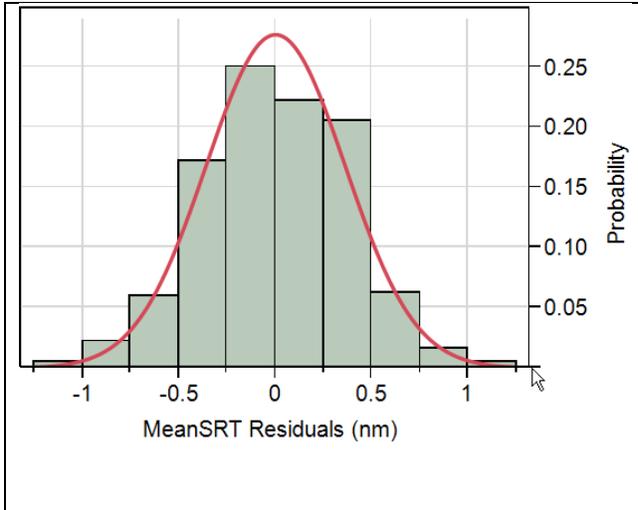


Figure 11. SRT Residual Histogram

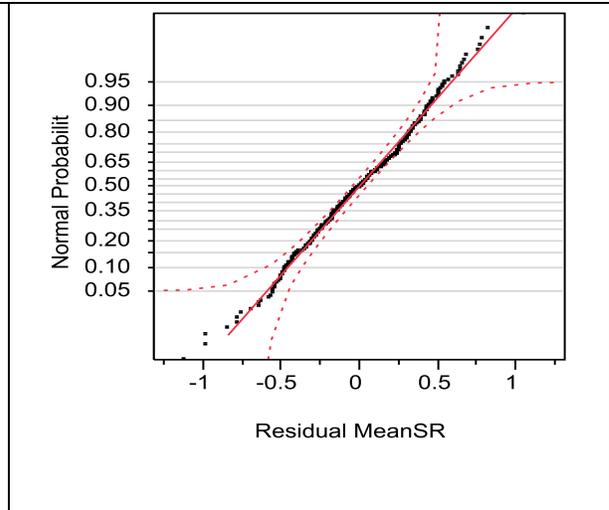


Figure 12. SRT Residual Normal Probability Plot

2. Vertical Adherence Parameter Calibration

The Vertical Adherence Algorithm has five parameters as illustrated logically in Figure 6 and defined below in Table 7. The parameters are again either continuous or fixed values. For vertical adherence, *changeStateToVertIn* and *changeStateToVertOut* are fixed parameters while *vertSpeedThreshold* and *timeWindowVert* are continuous parameters. The parameters *vertCnfThreshold1* and *vertCnfThreshold2* represent the range in feet per minute (fpm) that an aircraft has to be moving for it to be considered ascending or descending (not level). These values are a negative and positive value. To simplify the model, *vertCnfThreshold1* and *vertCnfThreshold2* are assumed to be equal with opposite polarities and referred to as *vertSpeedThreshold*.

Table 7. Vertical Adherence Algorithm Parameters

Parameter Name	Parameter Description	Related Model Values
<i>changeStateToVertIn</i>	These input properties specify the number of points that need to be determined to be vertically in or out, respectively, before the algorithm declares the point in or out, respectively.	1, 2, 3 points
<i>changeStateToVertOut</i>		1, 2, 3 points
<i>vertCnfThreshold1</i> <i>vertCnfThreshold2</i>	This input property is a range in feet per minute (fpm). This value is used to determine when an aircraft is ascending, descending, or level. One represents a negative value for descent while the other represents a positive value for ascent. For purpose of this experiment both are set equal and expressed as <i>vertSpeedThreshold</i> in the model.	<i>vertSpeedThreshold</i> 100.0 – 300.0 fpm
<i>timeWindowVert</i>	This input property, given in seconds, specifies an amount of time, correlating to a number of track points, to look back when determining the vertical speed of an aircraft at a given track time.	20 – 60 seconds

For the Vertical Adherence Algorithm, a full factorial experiment exercising all combinations of factor levels results in 36 (3 x 3 x 2 x 2) experimental runs plus two runs for testing non-linearity in continuous factors. Section III.0 defines the four response variables used to evaluate the different parameter settings.

The treatment effect defined in Eq. (3) and Eq. (4) is expanded to describe the model implemented for the Vertical Adherence Algorithm. Eq. (6) represents the treatment effect for the four factors in model design where the main and two-way interactions are specified. Ref. 9 includes a full explanation of the model and the included assumptions.[‡]

$$\tau = A + B + C + D + AB + AC + AD + BC + BD + CD + A' + A'B + A'C + A'D + B' + B'C + B'D \quad (6)$$

where *A* is the first factor, *B* is the second factor up to *D* which is the last factor, and *A'* and *B'* represent the third setting for the first and second factor

Figure 13 displays the leverage plots for two of the four response variables. The leverage plots illustrate a comparison between the actual and modeled values. The Rsq indicates how well the model was able to capture system variations. Similar to the lateral algorithm in the previous section, if the output variable has an Rsq of 1, it will be displayed with the actual versus modeled values plotted perfectly on the diagonal line of the leverage plot.

For this experiment, all four response metrics had high Rsq values, indicating that the majority of the system variation is being captured by the model. In Figure 13 leverage plots, the SRT response variable has an Rsq value of 1.00 and P(MC) has a value of 0.99.

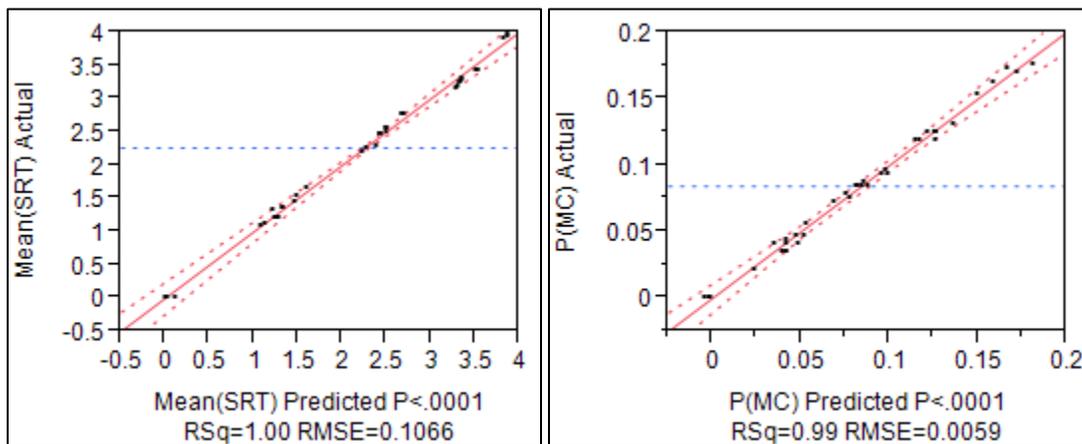


Figure 13. Vertical Adherence Leverage Plot for SRT and P(MC) Response Variables

Table 8 and Table 9 list the effect tests for the various factor level combinations of the experiment. They only provide results for one of the response variables, P(MC), but the other three look very similar. The p-values for all four response variables are listed in Table 12 of the Appendix. The definitions of the column headings displayed in Table 8 and Table 9 were previously described in Section III.C.1.

[‡] The non-linear factor was removed for the Vertical Algorithm’s model because it was proven to be insignificant.

Table 8. Vertical Adherence Model Main Effects for P(MC)

Source	DF	Sum of Squares	F Ratio	Prob > F
changeStateToVertIn	2	0.0012	17.2729	<.0001*
changeStateToVertOut	2	0.0585	846.8812	<.0001*
timeWindowVert	1	0.0108	297.3872	<.0001*
vertSpeedThreshold	1	0.0179	517.8437	<.0001*

Table 8 indicates that all four of the factors are statistically significant for P(MC) with p-values less than 0.05. For Table 9, the analysis continues for all the interaction effects which displays that some factor combinations are significant for P(MC) and others are not.

Table 9. Vertical Adherence Model Two Way Interactions for P(MC)

Source	DF	Sum of Squares	F Ratio	Prob > F
changeStateToVertIn*changeStateToVertOut	4	0.00010	0.67610	0.61770
changeStateToVertIn*timeWindowVert	2	0.00040	6.84890	0.00660*
changeStateToVertOut*timeWindowVert	2	0.00030	4.22750	0.03230*
changeStateToVertIn*vertSpeedThreshold	2	0.00110	16.23600	0.00010*
changeStateToVertOut*vertSpeedThreshold	2	0.00020	3.72820	0.04540*
timeWindowVert*vertSpeedThreshold	1	0.00010	4.02780	0.06090

Figure 14 displays the predictor profiler for the Vertical Adherence Algorithm. It can easily be determined by examining the magnitude of the slope that increasing the parameter *changeStateToVertOut* has a large impact on P(MC) and SRT, which is an expected result based on the idea that increasing the count would keep aircraft in adherence longer. It is interesting to note the amount of impact that changing the *timeWindowVert* has on all four of the output variables. As the amount of time increases all four output metrics increase as well, indicating that the determination of an aircraft flying level, ascending or descending is essential in the determination of vertical adherence. Furthermore, it is preferable to limit the window of track used to calculate this phase of flight determination.

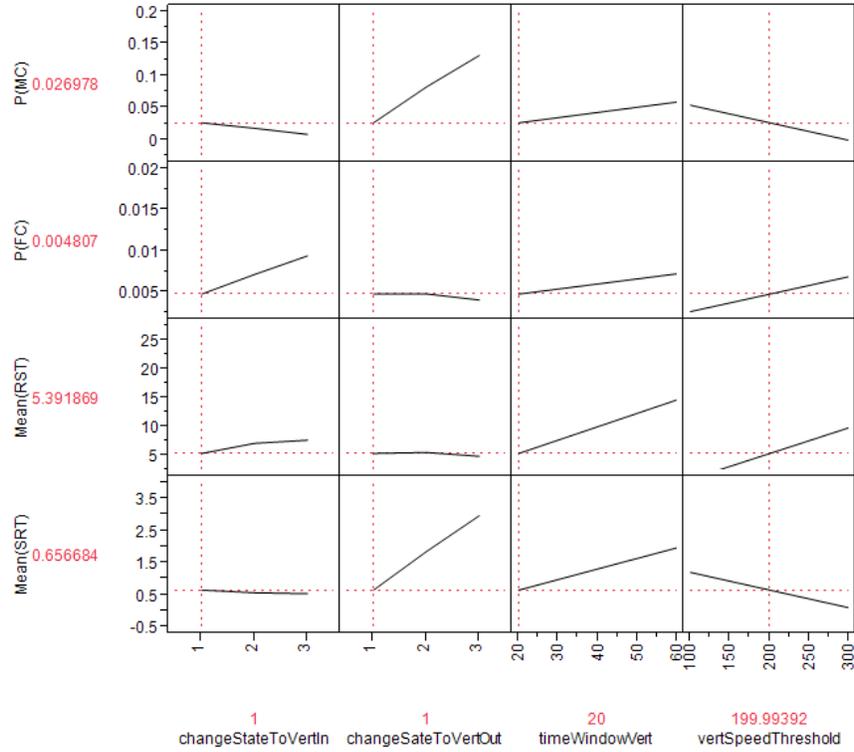


Figure 14. Vertical Adherence Model Results

The values of the factors listed on the bottom of Figure 14 are not the optimal settings but are close to optimal. They represent a combination of operational settings and optimal settings from the model with predicted values that produce good overall performance. To validate the model further, the parameter settings were actually run on the same test scenario and the results were compared to the model predictions. The model results are listed in Table 10. The resulting residuals are reasonably small.

Table 10. Model Results for Near Optimal Settings of Lateral Adherence Algorithm

	P(MC)	P(FC)	MeanRST (sec)	MeanSRT (sec)
Predicted	0.026978	0.004807	5.391869	0.656684
Actual	0.037618	0.004917	5.2	0.922222
Residual	-0.01064	-0.00011	-0.191869	-0.265538

To further confirm the model's validity, residuals for all the experimental runs were calculated and examined in detail. All four response variables produced distributions that were approximately normally distributed with means near zero. For example, the model's residuals of the P(FC) response variable is illustrated by histogram in Figure 15 and normal probability plot in Figure 16. Both provide a graphical diagnostic for testing if a data sample matches a normal distribution. In this case, the residuals or random error that the model captures is approximately normally distributed, has a mean close to zero, and a relatively small variation.

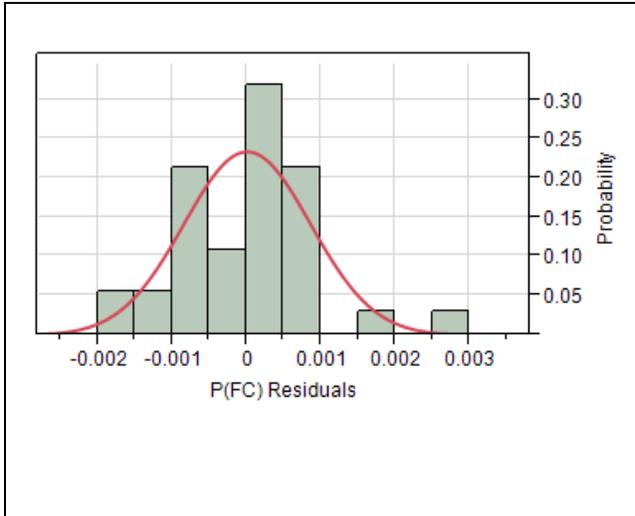


Figure 15. P(FC) Residuals Histogram

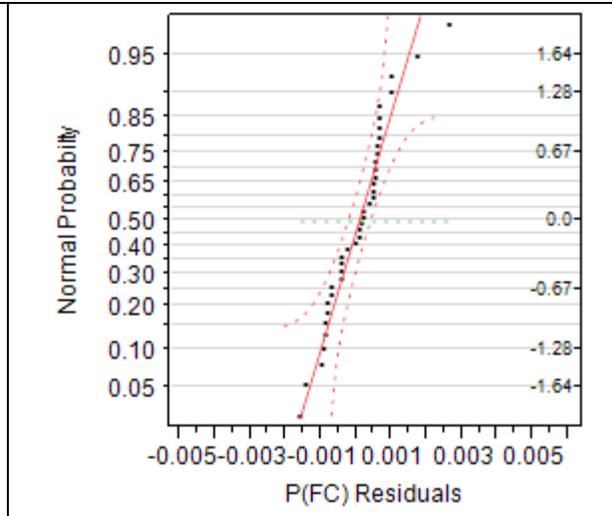


Figure 16. P(FC) Residual Normal Probable Plot

IV. Conclusion

To evaluate the trajectory and conflict predictions of NextGen decision support tools, the Concept Analysis Branch at the FAA utilizes two algorithms to determine if aircraft are following their known air traffic control clearances both laterally and vertically. The algorithms were documented and then evaluated under different parameter settings using experimental design techniques. To support this, a test bed was developed and a ground truth test data set defined. The resulting experiments helped calibrate the models by providing empirical evidence on which levels of each of the algorithm's various parameters would achieve near optimal performance. It was accomplished through application of advanced design of experiments and the use of an off-the-shelf statistical platform, JMP® statistical software⁸. Furthermore, the techniques presented in this paper can provide the readers with methods to develop and calibrate their own algorithms.

In summary, the FAA's Concept Analysis Branch was successful in developing, documenting, and calibrating a set of algorithms that allow filtering of its other DST performance metrics, like trajectory prediction accuracy. Thus, the methods documented allow the FAA to segregate laterally and vertically adhering flight data from situations where it is not adhering. Since DST functions are being advanced for NextGen, the FAA's Concept Analysis Branch can apply these techniques to properly evaluate the new functions and make evidence based recommendations to decision makers. This has already occurred in a number of cases. For example in 2011, it applied these techniques to measure the performance of NextGen funded advances in the En Route Automation Modernization (ERAM)¹⁰. It is anticipated these algorithms will continue to be utilized to support NextGen and other FAA initiatives for years to come.

⁸ JMP® is a commercial software package, for details on its full capabilities see <http://www.jmp.com/>.

Appendix

Appendix contains listing of statistical results of all modeled effects and output metrics. Table 11 contains results for lateral algorithm's experiment and Table 12 contains the results for the vertical algorithm's experiment.

Table 11. Full Statistical Results of Lateral Experiment

Source	Effect P-Value **			
	P(MC)	P(FC)	RST	SRT
innerLatThres	0.0000	0.0000	0.0000	0.0000
outerLatThres	0.0000	0.0000	0.0000	0.0000
changeStateToLatOut	0.0000	0.0000	0.0000	0.0000
changeStateToLatIn	0.0000	0.0000	0.0000	0.0000
bearingToNextFix	0.0000	0.0000	0.0000	0.0000
nextFixDistThres	0.0000	0.0000	0.0000	0.0000
longDistThres	0.0000	0.0000	0.0000	0.0000
innerLatThres*outerLatThres	0.7714	0.5854	0.0000	0.4318
innerLatThres*changeStateToLatOut	0.0017	0.0577	0.0000	0.2622
innerLatThres*changeStateToLatIn	0.2479	0.3644	0.0000	0.6769
innerLatThres*bearingToNextFix	0.0000	0.0000	0.0000	0.0000
innerLatThres*nextFixDistThres	0.5888	0.0087	0.0000	0.7486
innerLatThres*longDistThres	0.0000	0.0017	0.0083	0.0443
outerLatThres*changeStateToLatOut	0.0000	0.0000	0.0196	0.0000
outerLatThres*changeStateToLatIn	0.0000	0.0000	0.0000	0.0000
outerLatThres*bearingToNextFix	0.0000	0.0000	0.0000	0.0000
outerLatThres*nextFixDistThres	0.0000	0.0303	0.1609	0.0000
outerLatThres*longDistThres	0.0000	0.0205	0.0001	0.0000
changeStateToLatOut*changeStateToLatIn	0.0037	0.7678	0.8011	0.8479
changeStateToLatOut*bearingToNextFix	0.0000	0.0000	0.0000	0.0000
changeStateToLatOut*nextFixDistThres	0.8937	0.8576	0.5411	0.0330
changeStateToLatOut*longDistThres	0.5860	0.6255	0.7342	0.0006
changeStateToLatIn*bearingToNextFix	0.8652	0.0000	0.5455	0.0410
changeStateToLatIn*nextFixDistThres	0.9737	0.5543	0.9919	0.9938
changeStateToLatIn*longDistThres	0.8018	0.9237	0.9788	0.5310
bearingToNextFix*nextFixDistThres	0.0061	0.0000	0.0000	0.2858
bearingToNextFix*longDistThres	0.0002	0.0000	0.0000	0.0000
nextFixDistThres*longDistThres	0.0000	0.3784	0.1201	0.0000
outerLatThres*outerLatThres	0.0000	0.0000	0.0000	0.0000

** P-value is the smallest level of significance at which the null hypothesis would be rejected. If the p-value is less than the required α , the null hypothesis should be rejected. It is the same as "Prob>F" from Section III.C.1.

Table 12. Full Statistical Results of Vertical Experiment

Source	Effect P-Value			
	P(MC)	P(FC)	RST	SRT
changeStateToVertIn	0.0000	0.0000	0.5618	0.5320
changeSateToVertOut	0.0000	0.0000	0.0007	0.0000
timeWindowVert	0.0000	0.0311	0.0108	0.0000
vertSpeedThreshold	0.0000	0.0000	0.0002	0.0000
changeStateToVertIn*changeSateToVertOut	0.6177	0.3474	0.9929	0.7261
changeStateToVertIn*timeWindowVert	0.0066	0.7891	0.9962	0.9208
changeSateToVertOut*timeWindowVert	0.0323	0.0182	0.0045	0.0000
changeStateToVertIn*vertSpeedThreshold	0.0001	0.2161	0.9993	0.9486
changeSateToVertOut*vertSpeedThreshold	0.0454	0.0016	0.0027	0.0000
timeWindowVert*vertSpeedThreshold	0.0609	0.0618	0.0295	0.0006

References

- ¹Joint Planning and Development Office, *Concept of Operations for the Next Generation Air Transportation System, Version 2.0*, [online library], URL: http://www.jpdo.gov/library/NextGen_v2.0.pdf, June 2007.
- ²Paglione, M., Garcia-Avello, C., Swierstra, S., Vivona, R., and Green, S., “A Collaborative Approach to Trajectory Modeling Validation”, *24th Digital Avionics Systems Conference*, October, 2005.
- ³Paglione, M., and Oaks, R. D., “Implementation and Metrics for a Trajectory Prediction Validation Methodology,” *Proceedings of the American Institute of Aeronautics and Astronautics (AIAA) Guidance, Navigation, and Control Conference*, August, 2007.
- ⁴Mondoloni, S., Swierstra, S., and Paglione, M., “Assessing Trajectory Prediction Performance – Metrics Definition”, *24th Digital Avionics Systems Conference*, October 2005.
- ⁵Green, S. M., Vivona, R. A., Grace, M.P., and Fang, T. C., “Field Evaluation of Descent Advisor Trajectory Prediction Accuracy for En Route Clearance Advisories,” *American Institute of Aeronautics and Astronautics Guidance, Navigation, and Control Conference*, AIAA-98-4479, 1998.
- ⁶Brudnicki, D., Arthur, W., and Lindsay, K., “URET Scenario-based Functional Performance Requirements Document,” MTR98W0000044, MITRE CAASD, April 1998.
- ⁷Mondoloni, S., *Aircraft Trajectory Prediction Errors: Including a Summary of Error Sources and Data*, FAA EUROCONTROL Action Plan 16: Common Trajectory Prediction Capabilities [online library], URL: <http://acy.tc.faa.gov/cpat/tjm/>, July, 2006.
- ⁸Lindsay, K., “Results of a URET Operational Utility Experiment,” MTR99W0000081, MTRE CAASD, January 2000.
- ⁹Montgomery, D. C., *Design and Analysis of Experiments, Fourth Edition*, New York, NY, John Wiley & Sons, 1997.
- ¹⁰Crowell, A., et.al., “Evaluation of Parameter Adjustments to the En Route Automation Modernization's Conflict Probe,” DOT/FAA/TC-TN12/2, Federal Aviation Administration, December 2011.