Application of Generic Metrics to Assess the Accuracy of Strategic Conflict Probes

M. L. Cale, M. M. Paglione, H. F. Ryan, R. D. Oaks, J. S. Summerill
FAA William J. Hughes Technical Center,
Atlantic City International Airport, NJ 08405

Abstract

To support the goals of Free Flight, the FAA has sponsored the development of several ground-based conflict prediction tools, referred to as conflict probes, to aid the controller in the strategic planning of aircraft separation management. The developers of these tools have created performance metrics for their particular tool based on their system's design choices. This paper discusses a generic set of metrics that can be used to assess the accuracy of any conflict probe and presents the results of their application to the User Request Evaluation Tool (URET).

Introduction

In the United States the overall system of managing and controlling air traffic is known as the National Airspace System (NAS) and is administered by the Federal Aviation Administration (FAA). Surveillance radar provides aircraft position information to the ground controllers, radio navigation aids provide position information to the pilots, and very high frequency (VHF) radios provide voice communications between the aircraft and the ground. Detailed procedures involving restrictions on routing, speeds, and altitudes are an integral part of the NAS. These restrictions severely reduce the amount of aircraft traffic that NAS can accommodate, particularly when the weather is bad.

Free Flight is an air traffic control concept that will reduce the restrictions imposed by the NAS on the

aircraft flights, and thus increase the efficiency of aircraft operations, while maintaining safety. The goal of free flight is to provide "unrestricted opportunity for all to use the limited airspace in a manner that is efficient, effective, and equitable" [4].

To achieve the goals of Free Flight, broad categories of advances in ground and airborne automation are required. One of the most important ground based tools currently being developed is a conflict prediction tool, or conflict probe. A conflict probe is a decision support tool that will provide the air traffic controller with predictions of conflicts (i.e., loss of minimum separation between aircraft) for a parameter time (e.g., 20 to 40 minutes) into the future. At a minimum, a conflict probe predicts the flight path of an aircraft, continuously monitors that flight path from current aircraft position information, and probes for conflicts with other aircraft and incursions into restricted airspace. The tool also assists the controller in resolving the predicted conflicts and with alternative route planning in response to user requests. In contrast to the current, more tactical methods of air traffic control, a conflict probe supports Free Flight by aiding the controller in the strategic planning of aircraft separation management.

The FAA has sponsored the development of two prototype conflict prediction tools: the User Request Evaluation Tool (URET) developed by MITRE/CAASD and the Center-TRACON Automation System (CTAS) Conflict Prediction and Trial Planning (CPTP) tool developed by NASA Ames Research Center. The technical accuracy of

these tools is a critical issue to be addressed in planning for Free Flight Phase 1 and the future integration of these tools. NASA Ames and CAASD have created and applied performance metrics for their specific conflict prediction tools [5, 6]. The Traffic Flow Management branch (ACT-250) at the FAA William J. Hughes Technical Center (WJHTC) has defined a generic set of metrics that highlight the performance of any conflict probe: trajectory accuracy, conflict prediction accuracy, prediction stability and conflict notification timeliness [2, 3]. Since these metrics are independent of a particular system's design choices, they provide common measures to evaluate the performance of different systems. This paper discusses how these generic metrics can be used to assess the accuracy of any conflict probe, and presents the results of the application of the conflict prediction accuracy and conflict notification timeliness metrics to URET.

Conflict Probe Processing

A conflict probe is responsible for predicting into the future (e.g. 20 minutes) both the path the aircraft will fly, and potential conflicts the aircraft will have with other aircraft or with restricted airspace. As implemented in existing prototypes, the aircraft's trajectory and any conflict predictions are based on the flight information and track data received from the NAS Air Route Traffic Control Center (ARTCC) Host Computer System (HCS), weather forecasts from the National Weather Service, and detailed adaptation databases, including aircraft modeling information and system information relating to the airspace and procedures (see Figure 1). The conflict probe uses the most current flight intent and tracked position information received from the HCS to build and maintain an aircraft trajectory that predicts the flight path of the aircraft. The conflict prediction algorithms of the conflict probe then use these trajectories to determine future separation violations, e.g., to predict conflicts for notification to a controller. The conflict probe also provides the controller with a trial planning capability for use in determining resolutions to conflicts, and for responding to user requests.
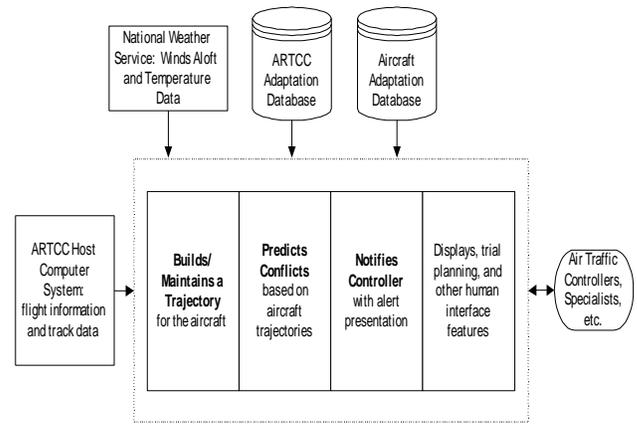

Figure 1: Components of the Conflict Probe Process

Measuring Conflict Probe Accuracy

The accuracy of a conflict probe is measured by its ability to correctly predict future events. This can be determined by measuring the difference between the conflict probe's prediction and what actually took place. That is, the accuracy will be a measure of the difference between what the conflict probe predicts to occur and presents to the controller, and what is determined to have actually occurred from post processing of the track messages from the HCS. The overall accuracy of the conflict probe predictions can be measured in several ways. ACT-250 has defined four broad categories of metrics: trajectory accuracy, conflict prediction accuracy, conflict notification timeliness, and prediction stability.

As previously mentioned, a conflict probe requires predicted aircraft trajectories to determine future separation violations, i.e., to predict conflicts. Thus, the **trajectory accuracy**, or the deviation between the predicted trajectory and the actual path of the aircraft, directly affects the accuracy of the conflict prediction. **Conflict prediction accuracy** is measured by several error probabilities that are used to quantify whether the predicted conflict actually occurred, and when a conflict that occurred was not predicted. In addition, the conflict predictions must not only be accurate in terms of the existence of a separation violation, but the conflict needs to be predicted in a timely manner. **Conflict notification timeliness** attempts to quantify the amount of lead-time the probe provides in the conflict predictions. Finally, the **prediction stability** metric quantifies the

stability of the various predictions made by the conflict probe. For example, a probe can make accurate trajectory or conflict predictions, but if these predictions change too frequently the user will have difficulty in making a choice between them. The following paragraphs describe these metrics in more detail.

Trajectory Accuracy

The trajectory prediction process is the foundation of any conflict probe, and as such, the accuracy of the predicted trajectory is important to the overall accuracy of the conflict probe. The path flown by an aircraft depends on the aircraft's flight plan, on its performance characteristics, on the weather, on the directions given by air traffic control, on the procedures required by the owner airline, and on individual pilot technique. A trajectory modeler takes into account all these factors, using both known and estimated data, to produce a prediction of the future aircraft flight path. Trajectory accuracy is the measure between the path of the aircraft predicted by the conflict probe and the aircraft's actual path. For our purposes, this is measured by the difference between the conflict probe-predicted trajectory and the tracked position reports received from the NAS HCS. This difference can be partitioned into two basic errors: spatial and time.

Spatial Error

The spatial error in the predicted aircraft position is measured by the slant range, or distance, between the predicted position (defined by a trajectory) and the actual position (defined by the HCS track reports) at a given time. A perfect prediction would have a slant range of zero. The slant range is divided into three orthogonal components: (1) lateral error, (2) longitudinal error, and (3) vertical error. The lateral error measures the side to side difference, or cross track error, between track and trajectory while the longitudinal deviation measures the along track distance difference between track and trajectory. The lateral and longitudinal errors comprise the horizontal component of the spatial error. These horizontal error components are illustrated in Figure 2 where the actual location of the aircraft (from the HCS track data) at time T is point A, the predicted location of the aircraft along the trajectory at time T is point B, and the predicted location of the aircraft along the

trajectory at the next point in time is point C. The distance between point B on the predicted trajectory and its corresponding track point A is the horizontal error; i.e., the total horizontal error is the length of the line AB. By drawing a perpendicular line from the point A to the line formed by points B and C, or its extension, an intersection point is determined, referred to as point D in Figure 2. The length of the line BD is defined to be the longitudinal error, and the length of the line AD is defined to be the lateral error.
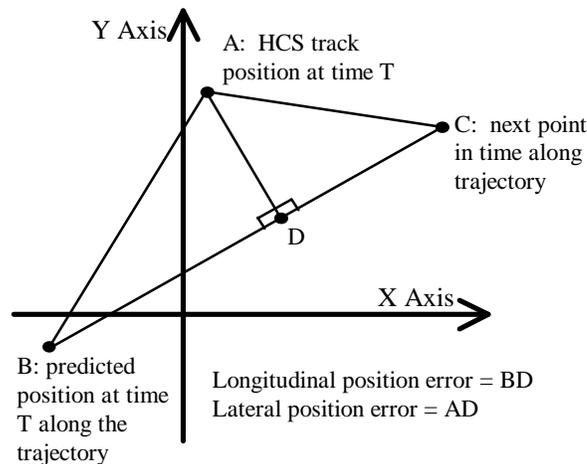

Figure 2: Definition of Longitudinal and Lateral Errors

The vertical component of the spatial error measures the altitude differential between track and trajectory. This is computed as the signed difference between the altitudes of two corresponding points from the HCS track data and the predicted trajectory data at the given time T.

Time Error

For the spatial error, a trajectory point and a track point at a coincident time are compared spatially. For the time error, a trajectory point and a track point at a spatially coincident location are compared in time. This error is a difference in the estimated time of arrival for a specific location, and is measured from the track point to the closest trajectory point in space.

Conflict Prediction Accuracy

The conflict prediction accuracy metric is the most significant metric category since it quantifies the

conflict probe's ability to meet its main goal: detecting conflicts. ACT-250 utilized a black box approach in defining the conflict prediction accuracy metrics (see Figure 3). Such an approach is only concerned with the input (i.e. the positions of the aircraft) and the output (i.e. predicted conflicts). Post-processing tools determine the actual conflicts using the aircraft position data, and compare these conflicts to the predicted conflicts.
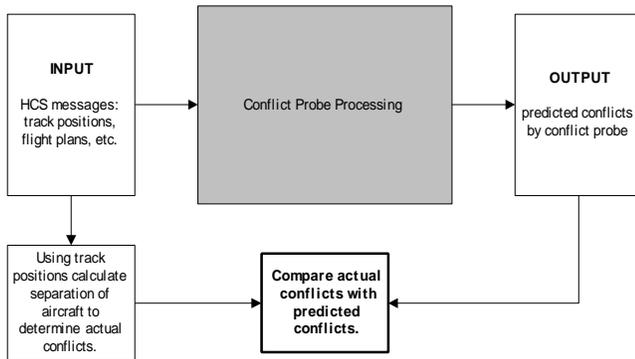

Figure 3: Conflict Prediction Accuracy Process

The conflict prediction accuracy metrics are based on two fundamental random events[1]: an aircraft **conflict** actually occurs, and an **alert** is presented to a controller based on a conflict prediction made by the conflict probe. The combination of outcomes of these two events is a random process related to the performance of the conflict probe. For example, the conflict probe may or may not predict a conflict between two aircraft, and in actuality, a conflict may or may not occur. Table 1 presents the four possible outcomes of these two random events: conflict and alert. For a Valid Alert (VA), a conflict occurs and the conflict probe correctly predicts an alert. For a False Alert (FA), an alert is predicted but a conflict does not actually occur, while for a Missed Alert (MA), a conflict occurs which the conflict probe does not predict. The fourth outcome includes all the remaining aircraft pairs for which no conflict

occurs and no conflict is predicted by the conflict probe.

	CONFLICT	NO CONFLICT
ALERT	**Valid Alert (VA)** probe predicts conflict / conflict occurs	**False Alert (FA)** probe predicts conflict / conflict does not occur
NO ALERT	**Missed Alert (MA)** probe does not predict conflict / conflict occurs	**Remaining Pairs (RP)** probe does not predict conflict / conflict does not occur

Table 1: Alert and Conflict Event Table

Figure 4 further depicts the possible outcomes of the random events presented in Table 1 as a Venn diagram. The correct prediction occurs for the two of the four outcomes represented by regions VA and RP in Figure 4, while the incorrect predictions are represented by the other two regions, FA and MA.
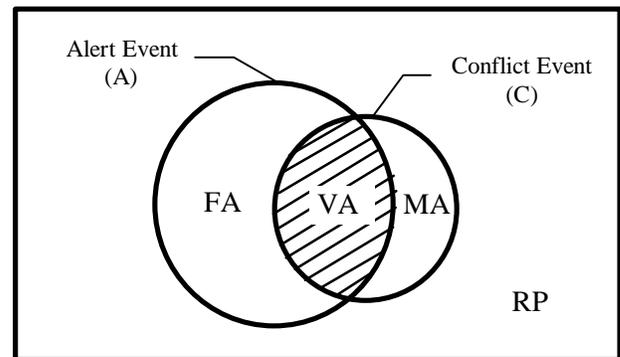

Figure 4: Venn Diagram of Alert and Conflict Events

The fundamental error events depicted in Figure 4, missed alert and false alert, can be modeled by conditional probability metrics that quantify the chances or likelihood of these outcomes occurring. It is necessary to use conditional probabilities to obtain meaningful results since, theoretically, any pair of aircraft in the airspace being probed can be in conflict, but most aircraft are widely separated from almost all of the other aircraft. Counting the number of occurrences of missed alerts and false alerts as a proportion of all the theoretically possible conflicts is thus not as meaningful as conditioning the probability on the conflict or alert events. For example, the probability of a false alert given an alert outcome

---

[1] Since there are many unknown factors that create an aircraft encounter, ACT-250 chose to model the encounter as a random event.

exists is a better measure of the performance, since it is not directly dependent on the size of the sample set.

The probability of a false alert, P(FA|A), is defined as the conditional probability that no conflict (C') exists given an alert (A) is predicted and presented to the controller. The estimate of the conditional false alert probability is simply the number of false alerts divided by the total number of alerts in the sample as expressed in the following equation.

$$P(FA|A) = \frac{P(C' \cap A)}{P(A)} = \frac{\text{probability of no conflict and an alert takes place}}{\text{probability of an alert}} \approx \frac{\text{number of false alert outcomes from sample}}{\text{number of alerts from sample}}$$

For the missed alert outcome, P(MA|C), the metric conditions on the conflict event and is expressed as the probability that no alert (A') is presented given an actual conflict (C) takes place. Similar to the method used for the false alert estimate, the missed alert estimate is the number of missed alert outcomes divided by the total number of conflicts in the sample (expressed in the following equation).

$$P(MA|C) = \frac{P(A' \cap C)}{P(C)} = \frac{\text{probability of no alert and a conflict takes place}}{\text{probability of a conflict}} \approx \frac{\text{number of missed alert outcomes from sample}}{\text{number of conflicts from sample}}$$

It is possible to make the missed alert probability as small as desired by changing a threshold value in the decision algorithm. However, when the probability of missed alert is decreased, the probability of false alert is increased. Similarly, the probability of false alert can be decreased at the expense of increasing the probability of missed alert. Therefore, both metrics have to be considered when examining the performance of a conflict prediction system.

Conflict Notification Timeliness

Conflict notification timeliness, measured in terms of the actual conflict warning time, attempts to quantify the amount of lead-time the probe provides to the controller for the conflict predictions. Conflict warning time is defined as the time difference between the time the alert is presented to the controller by the conflict probe and the actual time when the two aircraft are first determined to be in conflict (based on the HCS track data).

Related to conflict notification timeliness is the conflict start time delta. Conflict start time delta is the absolute value of the difference between the predicted start time of the conflict and the actual conflict start time. Thus it is a measure of the prediction error for the start of the conflict.

Prediction Stability

Simply stated, prediction stability is a measure of the frequency that the conflict probe "changes its mind" (i.e., creates new trajectories or switches between "conflict" and "no-conflict"). The stability of the trajectories can be measured by the number of different trajectories that are created for the aircraft in a given time period. For conflict predictions, stability will be measured by the number and rate of distinct alerts per real conflict situation, and per false conflict situation.

Application of Generic Metrics to the User Request Evaluation Tool (URET)

URET is an automated conflict prediction tool that is currently in use in prototype form as a decision support aid for the en route radar associate ("D-side") air traffic controllers at the Indianapolis and Memphis ARTCCs. URET detects aircraft-to-aircraft and aircraft-to-airspace conflicts for IFR aircraft tracked by the HCS, and provides alert information to the radar associate ("D-side") controller when such conflicts are detected. In February 1998, a simulation study was completed at the FAA WJHTC to determine the conflict prediction accuracy of the User Request Evaluation Tool (URET) Delivery 3 [1].

Approach

ACT-250's approach in accomplishing the URET conflict prediction accuracy study was to develop an Indianapolis simulation capability at the WJHTC. This approach was chosen, rather than using actual "real world" data, because one would expect there to be no conflicts in the actual data since any potential conflict would have been resolved by the controller prior to its occurrence. Use of a simulation allows the emulation of situations that can not be observed or completely controlled in the real world, and allows a wide range of test cases to be generated and analyzed using the basic infrastructure and analysis tools. Another advantage of this approach in

assessing the accuracy of conflict prediction tools is that the conflicts can be modeled at any minimum separation desired.

The WJHTC simulation testbed emulates the URET field installations with the NAS HCS connected to the URET system in the Traffic Flow Management (TFM) laboratory via the fielded URET HCS interface, the General Purpose Output Interface Unit (GPOIU). The simulation activity extracted flight plans from Indianapolis Center System Analysis Recording (SAR) tapes and used the Pseudo Aircraft System (PAS), a high fidelity aircraft simulator, to model the flights, without controllers separating the aircraft. This process produced a number of aircraft to aircraft conflicts. The simulated position data was transmitted to the WJHTC Target Generation Facility (TGF) which created simulated radar reports. The WJHTC HCS received the TGF radar messages, as well as flight plan information extracted from the SAR data, and provided flight plan and simulated track data to URET in the TFM laboratory via the GPOIU.

During the simulation run process, the HCS track reports provided to URET were recorded as were the alert records generated by URET. These two source files (i.e., what actually happened from the tracks and what was predicted from the alerts) were utilized during post-processing by the data reduction and analysis tool set to estimate the conflict prediction accuracy of URET. An overview of this complete process is depicted in Figure 5.

Estimation of URET Conflict Prediction Accuracy and Notification Timeliness

The conflict prediction accuracy and conflict notification timeliness metrics for aircraft to aircraft conflicts were estimated for URET Delivery 3 based on the results of the simulation study. These metrics, presented in terms of probabilities, quantify the likelihood of correct and incorrect conflict predictions. The underlying assumption of the metrics is that the HCS track data is the "ground truth" location of the aircraft.
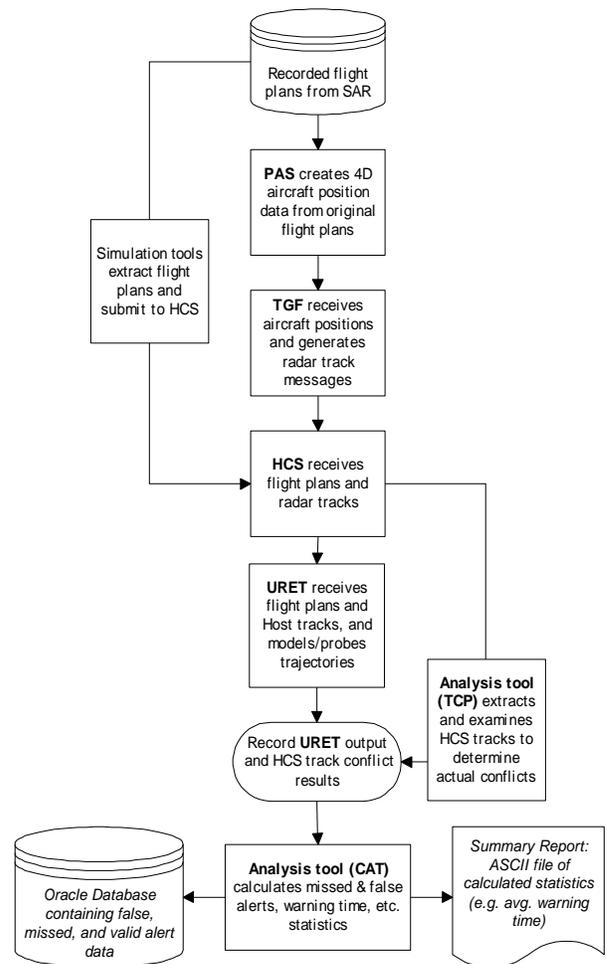

Figure 5: Simulation Overview

Conflict Prediction Accuracy

For this study, the accuracy of the conflict prediction is the measure of the difference between what URET predicts will occur (based on its predicted trajectory) and presents to the D-side controller as a URET alert, and what is determined to have actually occurred from post-processing of the HCS track messages. URET presents two levels of alerts to the D-side controller for predicted aircraft-to-aircraft conflicts: red alerts indicate that the centerlines of two URET aircraft trajectories are predicted to be in violation of standard separation, while yellow alerts

6

indicate that the conformance boxes[2] that surround the aircraft's trajectories are predicted to be in violation of separation standards. The following definitions apply:

- **Conflict** – A violation of separation standards between two aircraft is determined from the HCS track data during simulation post-processing.
- **Valid Alert** - An alert (red or yellow) is presented to a controller by URET and a conflict exists in the track data.
- **Late Valid Alert** - A conflict exists in the track data and URET presents an alert to the controller less than five minutes before the start of the conflict (at least five minutes of track data must exist for this to occur).
- **Missed Alert** - A conflict exists in the track data but URET did not present an alert to a controller.
- **Strategic Missed Alert** (SMA) - Expands on the missed alert definition to include the case when a conflict exists in the track data and URET presents an alert to a controller less than five minutes before the start of the conflict (i.e., a late valid alert).
- **False Alert** - An alert (either red or yellow) is presented to a controller by URET but a conflict does not actually exist in the track data.

Conflict Notification Timeliness

Conflict notification timeliness is measured in terms of the actual conflict warning time. The actual conflict warning time for each valid alert is the difference between the time the URET alert is presented to the controller (i.e. notified of an alert) and the actual time when the two aircraft are first determined to be in conflict (based on the track data).

As implied in the definition, the conflict warning time is calculated only for valid alerts and excludes the subset of valid alerts called late valid alerts. In addition to the late valid alerts, there are valid alerts with corresponding conflicts in our analysis that occur very early in recorded tracking of the flight (e.g., pop up situations). These valid alerts are excluded from the calculation of the warning time statistics since URET would not be able to provide sufficient warning time for the impending conflict situation. The current rule used for excluding such valid alerts consists of determining if the conflict start time for these alerts is less than a parameter time (e.g., five minutes) from the start of the track data.

Related to the conflict notification timeliness is the conflict start time delta statistic. Conflict start time delta is the absolute value of the difference between the predicted start time of the conflict and the actual conflict start time. Thus it is a measure of the prediction error for the start of the conflict. Like conflict warning time, conflict start time delta is calculated only for valid alerts that have sufficient track data from the start of the conflict (i.e., five minutes of track data) and does not include the subset of valid alerts called late valid alerts.

Data Collection, Reduction and Analysis

The effective estimation of the conflict probe accuracy metrics requires considerable data to be collected and analyzed. This data collection process was automated by the development of a generic data reduction and analysis (DR&A) tool set using an Oracle relational database system. The DR&A tools determine the difference between the actual track conflicts and the URET predicted conflicts, and compute the probability metrics; i.e., the accuracy of URET's conflict prediction. This tool set is comprised of the Track Conflict Probe (TCP) and the Conflict Analysis Tool (CAT), described in the following paragraphs.

Track Conflict Probe

The Track Conflict Probe (TCP) determines the actual conflicts that are in the scenario based on the "ground truth" HCS track data contained in the Oracle database. TCP first takes the raw track data, which is in twelve second increments, and time synchronizes it. This technique allows the analyst to interpolate through missing track points or to adjust the time to different size increments. The time

---

[2] URET centers a conformance box around the aircraft on its trajectory to represent regions of uncertainty. The nominal dimensions of the conformance box for aircraft in straight and level flight are +/-2.5 nautical miles laterally, +/-1.5 nautical miles longitudinally, and +/-300 feet vertically; the appropriate dimension(s) are expanded when an aircraft is turning or climbing/descending.

synchronized track reports are then stored in two Oracle tables, one which contains the actual time synchronized track reports and one which contains summary information for the track reports (e.g., maximum and minimum distances per dimension, etc.).

TCP looks at all $n(n-1)/2$ aircraft pairs (where n is the number of aircraft in the particular simulation run) to determine if there is some time overlap in the track data. Those aircraft with some time overlap are then subject to a gross filter which utilizes the track summary information contained in the Oracle database to determine whether horizontal and vertical separation requirements have been violated. Basically, this is accomplished by constructing three-dimensional boxes around the entire track set for each aircraft and computing the separation between the boxes. Aircraft pairs whose boxes are separated by less than 25 nautical miles in the horizontal dimension, and less than 5000 feet in the vertical dimension, are considered to "pass" the gross filter.

Only those aircraft pairs that pass the gross filter are checked for minimum separation. Separation violations are determined using an iterative process of comparing track co-ordinates at equivalent times and calculating separation distances between them. This process produces two additional Oracle tables describing the encounters of aircraft pairs in conflict and aircraft without conflicts. For both of these tables, two basic sets of separation distances and associated times are calculated for each pair of aircraft. The minimum horizontal separation (MHS) is the minimum distance in the horizontal dimension, measured in units of nautical miles. The minimum vertical separation (MVS) is the minimum distance in the vertical dimension, measured in units of feet.

Conflict Analysis Tool

The Conflict Analysis Tool (CAT) compares the output of TCP with the alerts notified to a controller by URET (contained in the Oracle database) and computes the following statistics:
- Probability of missed alert
- Probability of strategic missed alert
- Probability of false alert
- Conflict warning time
- Conflict start time delta

Basically, CAT examines each conflict determined by TCP and looks for a corresponding URET alert in the Oracle database. Only unique, notified URET alerts are counted as alerts (i.e., new alerts as opposed to updates to existing alerts). Those conflicts that have no corresponding URET alert are considered Missed Alerts. Alerts which have a corresponding conflict identified are flagged as Valid Alerts; duplicate alerts are excluded from the analysis. The subset of Valid Alerts determined to have conflict warning times less than five minutes are considered Late Valid Alerts. By process of elimination, all remaining alerts for which there is no corresponding conflict are initially considered to be False Alerts. The false alerts are then filtered for exceptions based on the following set of rules:
- URET alerts that predict a conflict outside the available track data for either involved aircraft are eliminated (i.e., the URET predicted conflict begins or ends before/after the existence of "ground truth" track data).
- URET alerts for aircraft that never had any track data (and therefore no "ground truth" data for TCP to use to determine a conflict) are excluded from analysis.
- URET alerts posted beyond the end of the scenario run, or the last track report time, are eliminated.

Results and Analysis

This paper presents the results of an analysis of the technical accuracy of URET conducted at the FAA William J. Hughes Technical Center. Nine simulations were conducted, each around 5 hours in duration with approximately 400 to 500 simulated aircraft. These nine simulations were analyzed twice with two different definitions of aircraft to aircraft conflicts. The first analysis, referred to as Analysis A, used basic standard radar separation for en route airspace as defined in FAA Order 7110.65, 4-5-1.a/b and 5-5-3.b.1 (i.e., five nautical miles in the horizontal dimension, and 1000 feet at or below flight level 290, or 2000 feet above flight level 290, in the vertical dimension). This represents the true standard separated conflict situation. The second analysis, referred to as Analysis B, expanded the separation distance in the horizontal dimension to ten nautical miles, which more closely models the encounter distances URET uses in its predictions of yellow

alerts for aircraft to aircraft conflicts. The results of these two analyses are summarized in Table 2.

	ANALYSIS A	ANALYSIS B
CONFLICT PREDICTION	ACCURACY	
MIssed Alert (MA) Probability	0.002	0.009
Strategic MA Probability	0.04	0.07
False Alert Probability	0.85	0.65
AVERAGE CONFLICT WARNING TIME	15 minutes	13.5 minutes
CONFLICT START TIME DELTA	3.7 minutes	2.2 minutes

Table 2: URET Conflict Prediction Analysis Results

Reported missed alerts based on the two analyses had a very low probability.   One (1) missed alert was detected in Analysis A, yielding an average missed alert probability of 0.002 for the nine simulation runs. Thirteen (13) missed alerts were found in Analysis B which increased the probability to 0.009.   The strategic missed alert probability is higher at 0.04 for Analysis A, and 0.07 for Analysis B. This is to be expected since the strategic missed alert includes missed alerts based not just on their presence but on their timeliness of being presented.

The false alerts were much more common and possibly a reason for the low missed alert probability, since these two fundamental errors have an inversely proportional relationship (i.e., a false alert probability that is high will provide a low missed alert probability).   For the nine simulation runs processed for Analysis A, on average, the conditional probability that a given URET alert does not have a matching actual standard separation violation is 0.85. When the actual conflict definition was increased to ten nautical miles for Analysis B, this number was reduced by one fourth to 0.65.

The false alert probability can also be conditioned on the specific URET alert level (i.e., red or yellow). When the probabilities are conditioned in this manner for Analysis A, the conditional probability for a red alert being false is 0.72, and 0.93 for yellow. In other words, given a red alert, the probability that the alert is false is 0.72 for Analysis A.   Similarly, for Analysis B, the conditional probability for a red alert being false is 0.51, and 0.73 for yellow.

The average warning time is estimated at 15 minutes and the average conflict start time delta is 3.7 minutes for Analysis A.   There is a small reduction in these times for Analysis B, where the average warning time is estimated at 13.5 minutes and the average conflict start time delta is 2.2 minutes.   This decrease is expected since an aircraft pair will be in violation of a ten nautical mile separation earlier than for a five nautical mile separation.

The missed and false alerts can be influenced by many factors, e.g., actual aircraft separation, encounter angles between aircraft, altitudes during aircraft encounters, aircraft types in the encounters, and warning time.   These factors may characterize the strengths and weaknesses of a particular conflict prediction approach as they may affect the performance of each conflict probe differently.   For example, Figure 6 presents the percentage of false alerts determined in Analysis A as a function of the minimum horizontal separation measured from HCS track reports. This distribution has a mean of around 9.5 nautical miles. Considering all the alerts without regard to alert level (i.e., red or yellow), the measured performance is consistent with URET's

Figure 6: Percentage of False Alerts Versus
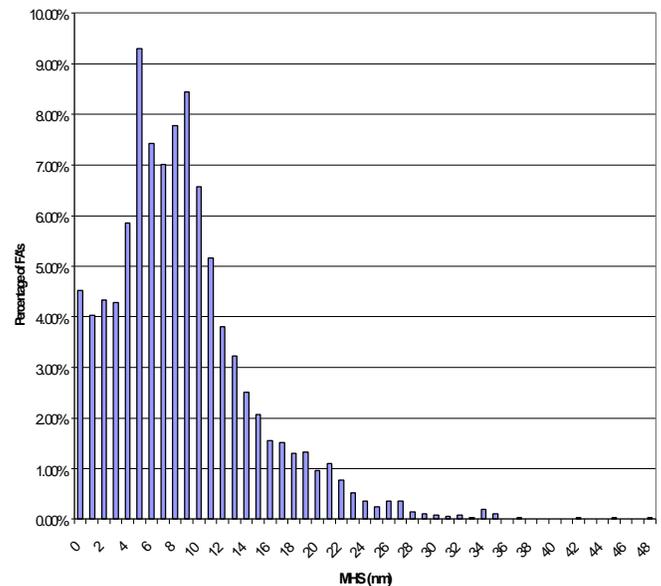
Minimum Horizontal Separation (MHS)

design since, depending on the encounter geometry, it predicts conflicts that are roughly between 8 and 15 nautical miles, with a nominal value of around 10 nautical miles. For Analysis B, there was a shift in the distribution making the mean minimum horizontal separation around 11 nautical miles. In other words, if the user considered a conflict at 10 nautical miles, URET still had some false alerts, and they were on average between aircraft with a MHS of about 11 nautical miles.

Conclusion

In quantifying the accuracy of a conflict probe, two fundamental errors are present: missed alerts and false alerts. A conflict probe is designed to meet acceptable limits of both these errors and a balance between the two is required. In support of Free Flight efforts and future integration needs, URET's conflict prediction accuracy and conflict notification timeliness were measured based on multiple Indianapolis simulation runs of URET. The end result is a set of descriptive generic statistics that estimate the general technical accuracy of URET's predictions. However, while these results are useful in defining functional performance requirements and establishing baseline accuracy levels for future URET deployments (e.g., URET Core Capability Limited Deployment (CCLD) for Free Flight Phase 1), they should not be used to make operational suitability judgements. That is, while the results provided in this paper give an indication of the technical performance of URET (i.e., "how good it is"), additional research and human factors evaluations are required to determine the technical performance required in an operational context (i.e., "how good it has to be").

ACT-250 is currently preparing to conduct additional analyses, using actual field data adjusted to include predefined conflict situations, that will expand this conflict prediction accuracy estimation to include the application of the complete set of generic metrics described in this paper to both URET and CTAS. In addition, future detailed designed experiments are planned to maximize the inferences provided by the data and determine the statistical influence of selected factors (e.g., encounter geometry, flight path characteristics, etc.) on the metrics. Future human factors evaluations will be conducted by other FAA organizations to determine the required technical accuracy for an operationally suitable conflict prediction system.

Acronyms

ACT-250	WJHTC TFM Branch
ARTCC	Air Route Traffic Control Center
CAASD	Center for Advanced Aviation System Development
CAT	Conflict Analysis Tool
CTAS	Center-TRACON Automation System
DR&A	Data Reduction and Analysis
FA	False Alert
GPIOU	General Purpose Output Interface Unit
HCS	Host Computer System
MA	Missed Alert
NAS	National Airspace System
NASA	National Aviation and Space Administration
SAR	System Analysis Recording
SMA	Strategic Missed Alert
TCP	Track Conflict Probe
TFM	Traffic Flow Management
TGF	Target Generation Facility
TRACON	Terminal Radar Approach Control
URET	User Request Evaluation Tool
VA	Valid Alert
WJHTC	William J. Hughes Technical Center

References

[1] *URET Conflict Prediction Accuracy Report* (DOT/FAA/CT-TN98/8), Cale, M., et.al., WJHTC/ACT-250, April 1998

[2] *Generic Metrics and Statistics to Estimate the Conflict Prediction Accuracy of Conflict Probe Tools* (draft), WJHTC/ACT-250, August 1997

[3] *Generic Metrics and Statistics to Estimate Trajectory Accuracy* (draft), WJHTC/ACT-250, October 1998

[4] *Free Flight Action Plan*, RTCA Incorporated, August 1996

[5] *A Methodology for the Performance Evaluation of a Conflict Probe* (AIAA-98-4238), K. Bilimoria, NASA Ames Research Center, August 1998

[6] *URET Scenario-based Functional Performance Requirements Document* (MTR98W0000044), D. Brudnicki et. al., MITRE/CAASD, April 1998