Lateral Intent Error's Impact on Aircraft Prediction

Mike Paglione, Ibrahim Bayraktutar, Greg McDonald, and Jesper Bronsvoort

The lateral deviation between the air traffic automation's known horizontal route of flight and the actual aircraft position is due to the typical navigation and surveillance errors, as well as the larger atypical errors that are mainly caused by purposeful changes in the route of flight that are not updated. The paper presents large data analyses of the ground automation systems of the United States and Europe, indicating errors from 20 to 30 nautical miles are common, while airborne Australian and more samples in the United States had errors from 100 to 800 times smaller. Further analysis illustrated the direct impact these errors have on safety critical separation management functions. It was concluded that airborne derived data via Automatic Dependent Surveillance Contract reports offer a major opportunity to improve the ground-based automation functions.

INTRODUCTION

Despite the current economic slow down, 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

---

Mike Paglione is with the Federal Aviation Administration, Atlantic City, United States. Ibrahim Bayraktutar is with EUROCONTROL, Brussels, Belgium. Greg McDonald and Jesper Bronsvoort are with Airservices Australia, Melbourne, Australia.

Received October 8, 2009; accepted January 4, 2010.

Air Traffic Control Quarterly, Vol. 18(1) 29–62 (2010)
 CCC 1064-3818/95/030163-20

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.

The interagency Joint Development Planning Office (JPDO) in the United States foresees a traffic demand increase by 2025 up to three times the number of flights of today's traffic [1]. 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 2025 [1]. 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.

Beginning in July 2004, the European Commission established a consortium of air traffic stakeholders with similar objectives for Europe, known as the Single European Sky Air Traffic Management (ATM) Research Initiative (SESAR). SESAR requires development of technology, standards, and procedures over the next eight years. The overall objective is to increase air traffic capacity by three while cutting aviation costs in half, improving safety by a factor of ten, and reducing the environmental impact of each flight by ten percent [2].

While the initiatives in Europe and the United States were still just discussions among aviation stakeholders, Australia embarked on a world first initiative to develop an ATM Strategic Plan as early as 1999. Based on a collaborative approach with User Preferred Trajectories as the ultimate goal; the ATM Strategic Plan establishes a framework that enables Australia to keep at the forefront of the Communications, Navigation, and Surveillance (CNS) systems and ATM development and its associated benefits.

The successful achievement of these ambitious initiatives set forth by multiple nations will require researchers across the globe to question their old paradigms within the existing processes and infrastructure to develop new approaches for meeting the challenges in these plans. With such high goals, there will be increasing demands in schedule and cost (doing more with less) so collaboration is vital to leverage resources and expertise.

This paper brings researchers together from the United States, Europe, and Australia to examine one specific, yet critical component within the aviation system–the understanding of the impact of lateral intent information within our current ATM automation and how it may improve in the future. The lateral intent of an aircraft is indeed only one aspect of the trajectory input information required to predict an aircraft's future path, but the challenge involved is a common problem across the globe. It is also of critical importance to the feasibility of automation that will assist in the separation management of aircraft. Therefore, global collaboration on the issues and potential solutions is warranted. This paper will first describe the trajectory prediction process followed by detailed explanations of the problem of missing lateral intent. Data and analysis results are presented in the next two sections that illustrate the magnitude of the problem. In closure, potential solutions are proposed.

AIRCRAFT TRAJECTORY PREDICTION PROCESS

Many of the operational concepts among the JPDO, SESAR, and Australia's ATM Strategic Plan promote 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 of the airspace problems 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, allowing air traffic supervisors to make more efficient adjustments to either avoid the condition or alter staff and/or 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.

Adapted from Reference 3, Figure 1 illustrates an example of the trajectory prediction process as applied to a commercial flight already en route. This example refers to a generic ground-based trajectory prediction and generation process; some TPs may require more, different or less information. The notional trajectory presented illustrates a simplistic trajectory prediction. A more

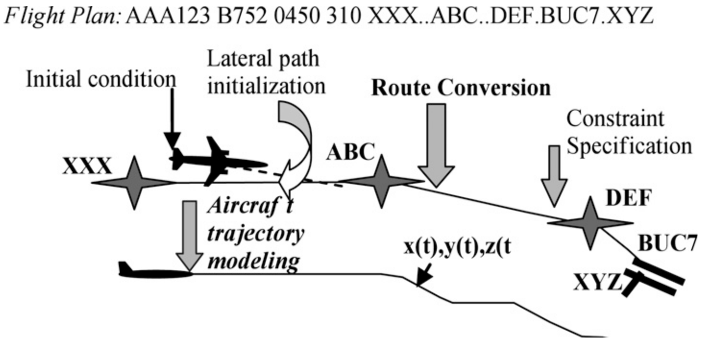

Figure 1. Trajectory generation – Adapted from [3].

complex instantiation of this process could lead to the inclusion of more sophisticated steps such as aircraft behavior modeling, in-flight parameter estimation, or trajectory error monitoring. The TP requires access to the flight plan containing the flight number (e.g., AAA123), the aircraft type (B-757-200), the filed cruise speed (true airspeed of 450 knots), the desired cruise altitude (31,000 feet), and the route of flight (from waypoint XXX, direct to ABC, then DEF, finally to XYZ via the BUC 7 STAR). Furthermore, the TP will have an estimate of the initial condition (present aircraft position, altitude, ground speed and ground track). Prior to conducting trajectory prediction, the flight plan route, expressed as named waypoints, jet routes, STARs, etc. will be converted to a series of geographical points (e.g., latitude and longitude). This process is known as *route conversion* [3].

Once the route is converted, a mechanism for joining the initial condition to the converted route is required. This process is referred to as *lateral path initialization*. This process may simply involve the identification of the initial location on the route. At times, the initial condition will be slightly off-route and some connection from the initial condition to the route will be required. A more generalized form of this trajectory service includes lateral intent modeling, in which the lateral path is generated based on assumed operational procedures of the pilot and/or controller.

Once the lateral path is determined, vertical and speed constraints must be considered at different points along the route of flight. This is the process of *constraint specification*. For example, speed constraints below 10,000 feet can be applied, as can altitude constraints along a standard terminal arrival route. The concept of *longitudinal intent modeling*, while implicit in some TPs, refers to the addition of speed and altitude procedural considerations that reflect how the combined controller, pilot and aircraft guidance system will "fly" the aircraft. An example is the estimation of the top-of-descent location or the planned descent speed.

All of the above steps are typically conducted prior to the calculation of a trajectory using physics-based modeling. We refer to this collection of information as the *preparation process*. The next step is the core or computation part of aircraft trajectory prediction. During this step, the lateral and vertical path are merged to “reflect” the predicted behavior identified in the preparation process, including: following the converted route, meeting specified constraints (such as altitude and speed constraints), following appropriate aircraft dynamics (such as turns, climbs and descents), and reflecting environmental and aircraft-specific effects. The output of this process is the predicted evolution of the aircraft’s state as a function of time.

PROBLEM OF MISSING LATERAL INTENT

The accuracy of the TP process described above 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 TP has a direct and significant impact on the DST’s overall performance and usability.

In addition to those previously described, the TP requires many inputs to produce an accurate trajectory prediction such as aircraft model characteristics, surveillance position reports, wind and temperature forecasts, and flight path intent information to name a few [4]. These factors have been the subject of many scientific studies. In [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 [6], researchers at the MITRE Corporation developed models to evaluate their DST’s overall performance by utilizing accuracy statistics of their TP’s performance. In yet another effort [7], 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 accuracy.

Under present-day operations, as illustrated previously in Figure 1, the flight plan message is the typical means of coding the aircraft operator’s request and air traffic control’s clearance of the aircraft’s horizontal 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 currently not often processed the same by the 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 horizontal 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 ATM automation [8].

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 in March 2005 of a civilian airliner traveling through the United States' Washington Air Route Traffic Control Center (ARTCC), referred to as ZDC. It originated from Dallas Fort Worth, Texas with the destination of John F. Kennedy International Airport (JFK) in New York. Figure 2 illustrates the top down stereographic view of the aircraft's horizontal path overlaid on the ZDC high-altitude sectors, which it travels through. On its journey to JFK, the sample flight is traveling in a northeasterly direction where ZDC accepts air traffic control for the flight at 20:14 UTC (Coordinated Universal Time) from adjacent Indianapolis ARTCC.

The focus of this example is the ground automation's trajectory built at 74005 seconds (20:33:25 UTC). This trajectory is illustrated in both Figure 2 and Figure 3 (darker segmented line) and overlaid with the surveillance track positions (light gray thicker line). Of particular interest is the complete hold maneuver performed later in the flight beginning roughly at 20:50 UTC. 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. An extraction of the trajectory metrics calculated for the 74005 second trajectory is listed in Table 1. A sample was taken at 74040 seconds (20:34:00) with a look-ahead time every five minutes up to 20 minutes in the future. At the first measurement time at look-ahead time of zero, the horizontal error (i.e. straight-line unsigned

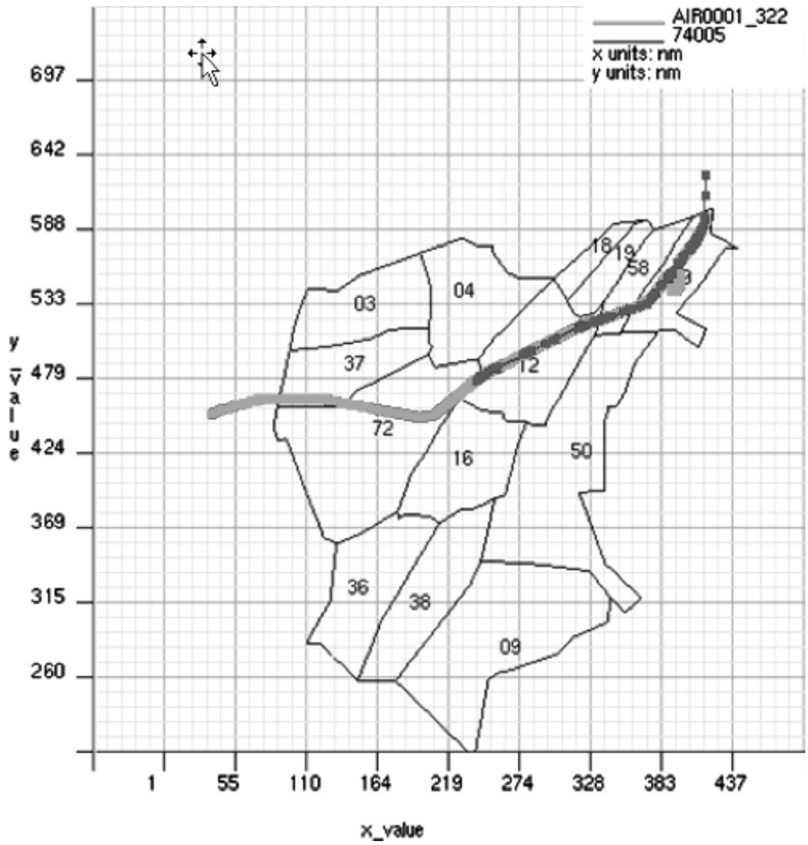

Figure 2. Sample flight, top down view.

error) was nearly half a mile with zero vertical error. However, as the look-ahead time progresses and approaches the turn as depicted in close-up view in Figure 3, the horizontal errors increased significantly. Due to the missed maneuver, the error reaches up to 32 nautical miles horizontally. The clearly visible cross-track error (i.e. side-to-side lateral error) is approximately 12 nautical miles, but the bulk of the error is found in the along-track error (i.e. longitudinal or along the route error). The additional travel time caused by the hold maneuver manifests in a -32 nautical mile along-track error, which translates to as much as 4.4 minutes lag in the trajectory prediction.

Clearly, if such lateral maneuvers in the form of heading vectors, holds, or changes in the horizontal path of an aircraft are not known by the ground based TP, they can cause large errors in trajectory predictions as shown by this example in Table 1. In the next section, metrics will be defined and the results of a large data analysis effort will further illustrate the magnitude of these lateral errors throughout the ATM system today.

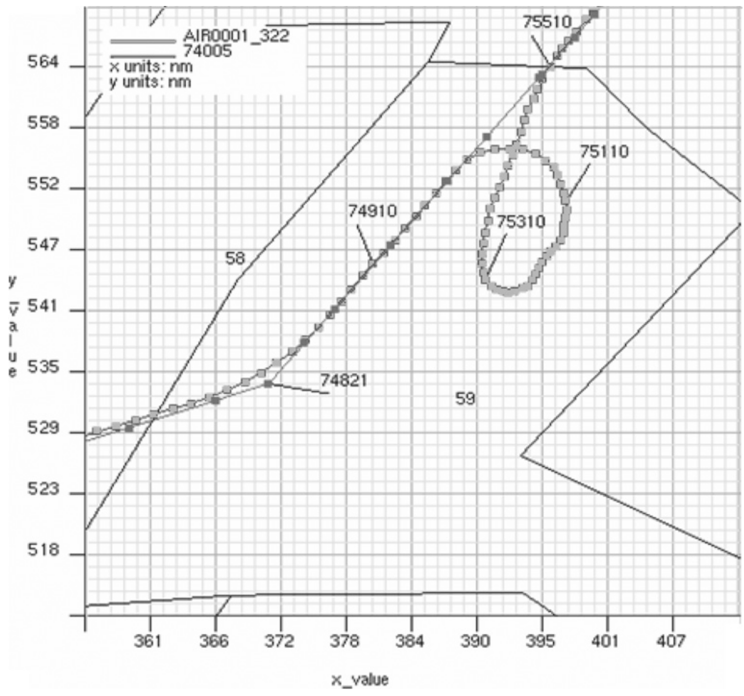

Figure 3. Close-up view of actual versus trajectory X-Y plot.

Table 1. Sample Flight's Trajectory Metrics

Measurement Time	Look-Ahead Time	Horizontal Error	Cross-track Error	Along-track Error	Vertical Error
HH:MM:SS	Seconds	Nautical Miles	Nautical Miles	Nautical Miles	Feet
20:34:00	0	0.4	0.3	-0.3	0
20:39:00	300	0.1	-0.1	0.0	793
20:44:00	600	1.2	-0.5	-1.0	0
20:49:00	900	2.1	-0.1	2.1	2096
20:54:00	1200	34.6	11.9	-32.5	6952

LATERAL DEVIATION METRICS

As described in the previous section, missing lateral intent data are often the result of various maneuvers being initiated without proper updates to the TP (typically located in the ground automation system utilized by DSTs). This error can be detected in post processing or even operationally by measuring the difference between the automation's known horizontal position and the coincident surveillance position. In a study conducted in [9] and in another in [10], the overall adherence to the current air traffic control clearance is defined as the status of whether the aircraft is following its known clearance at each instance of time during its flight. As with any definition, it is subject to interpretation, but focusing only on the lateral dimension discussed in [10], it is interpreted to mean that the surveillance radar position (or global positioning satellite position if available) for an aircraft should be declared out of lateral adherence when it is determined that the aircraft's intent was to deviate laterally from its known cleared route.

Figure 4 shows the geometry associated with determining lateral adherence of a surveillance data point for an aircraft. The figure shows an aircraft at a specific position and flying along a path in a specific direction. The figure also shows the current route segment for the aircraft with a triangle depicting the next fix on this route. Identified in this figure are two key parameters that can be used to define whether or not an aircraft is in lateral adherence at a specific position. They are:

1. d_r - The normal distance from the aircraft's actual position point to the aircraft's current route segment.
2. β–The angle between the aircraft's direction of flight and a line drawn from the aircraft's actual position point to the next fix. (This will be referred to as bearing to the next fix in this paper although an aircraft's direction of flight may not be equal to its heading.)

Based on the geometry depicted in Figure 4, the actual position of an aircraft would be considered to be in "perfect" lateral adherence when both the β angle and d_r distance are equal to zero. Operationally, aircraft rarely, if ever, exhibit such behavior. Therefore, if these values are within certain predetermined thresholds, it could be stated that the aircraft is in lateral adherence to the current known route.

Even though these metrics and their measured distributions are universal, the combinations of thresholds chosen to determine if an aircraft is in a state of lateral adherence are truly dependent on the DST application being supported. For this paper the thresholds utilized were developed originally in [11] to support a non-operational conflict probe. Furthermore, a heuristic method was

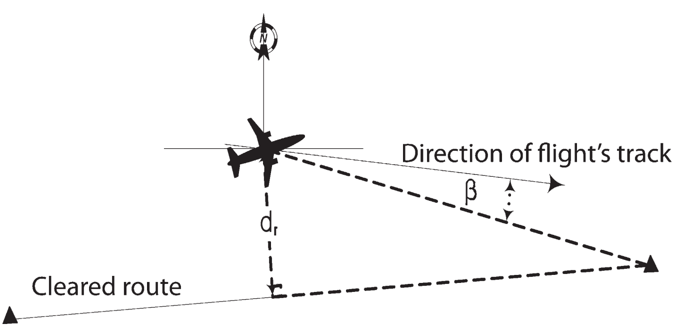

Figure 4. Geometry of lateral deviation – Adapted from [10].

implemented in [11] to determine if the DST's TP should utilize the flight plan or base its prediction strictly on course heading information from radar surveillance data. This heuristic approach provides descriptive states of lateral adherence and non-adherence that can be utilized to quantify operational data.

The heuristic algorithm is illustrated in the following Figure 5. It begins by calculating the normal distance to the route, d_r, (or simply the lateral deviation) and the angle β to the next fix position on the route. If at the end of the route, immediately end the calculation and label the state, *endOfRoute*. This is an indeterminate case when the

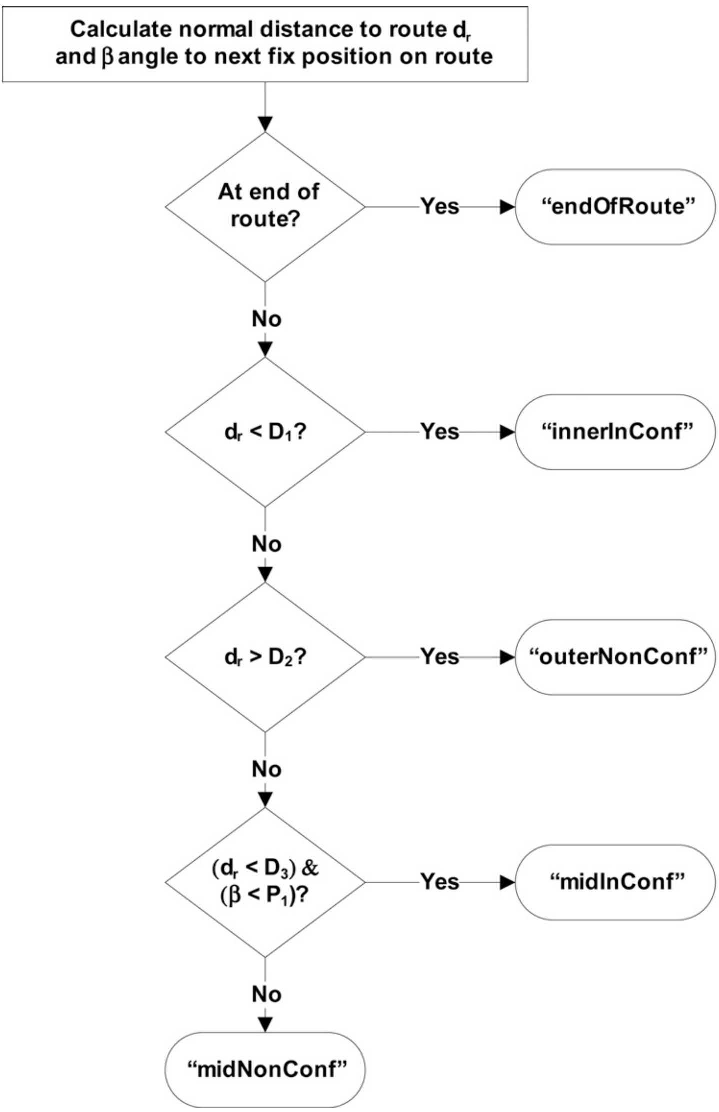

Figure 5. Lateral adherence heuristic states.

aircraft has gone past the end of the flight plan route and thus lateral intent is unknown. If not at the end of the route, the lateral deviation is checked and if below a threshold, D_1, is labeled to be in a state of *innerInConf*.

If the lateral deviation is larger than the threshold D_1, then it is checked against a larger threshold D_2. If it exceeds this threshold, it is labeled *outerNonConf* for the first out of lateral conformance state. If the lateral deviation is below, both the lateral deviation and angle β are checked against thresholds D_3 and P_1, respectively. If both measures are below their respective threshold, the state is labeled *midInConf* for the last in conformance state or not as *midNonConf* as the last out of conformance state. The thresholds used in this paper are listed in Table 2.

LATERAL INTENT RESULTS ON OPERATIONAL DATA

The late professor and modern management theory pioneer, Peter F. Drucker, is attributed with the famous quote, "If you can't measure it, you can't manage it." Following this advice in perspective of the TP, intent, and deviation information subsequently presented, the previously described metrics are applied to a large set of operational data from the United States, Australia, and Europe. This section provides a guide to the magnitude of the error in various contexts on the ground and in the air. For the ground automation data, it also includes an analysis of aircraft conflict predictions.

Lateral Deviation Results for Ground-Based Automation

This paper utilizes United States air traffic data collected for a recent study in [11]. It includes seven hours of air traffic messages, amounting to approximately 50,000 flights, and associated adaptation (i.e. detailed definitions of airspace boundaries and fix locations for expanding the flight plans) collected on April 3, 2008 for all twenty en route ARTCCs within the continental United States. The air traffic messages were retrieved from the Host Air Traffic Management Data Distribution System (HADDS). The messages record each ARTCC's air traffic control clearances and surveillance radar track positions.

Table 2. Thresholds for Lateral Adherence States

Threshold	Value (units)
D_1	0.5 (nm)
D_2	1.5 (nm)
D_3	1.0 (nm)
P_1	30 (deg)

Next, the messages are parsed and planned routes are expanded from the flight plan amendments into a series of geographic positions. The data sample captures the afternoon peak traffic schedule including the traffic messages recorded from 17:00:00 to 23:59:59 UTC (Coordinated Universal Time). The selected expanded routes and associated surveillance radar positions are the input data source for generating the lateral intent error metrics defined earlier. Table 3 provides the listing of ARTCC code versus location and the flight count per sample.

For the European data discussed in this paper, the statistics are cited from the EUROCONTROL Flight Data Management Metrics project, published in [12]. The project's objective is to measure the quality of flight data available to stakeholders, including data consistency, accuracy and other measures. The data set represents a large European flight sample collected for one day in November 2006. Approximately 27,000 flights from EUROCONTROL's Central Flow Management Unit were supplied by 31 European Air Navigation Service Providers (ANSPs) across Europe [12].

Lateral Deviation Statistics. As cited in [12], the European data collection was conducted with the aid of the EUROCONTROL Flight Information Consistency Analysis Tool (EFICAT). For the two-dimensional route analysis in [12], the field data was grouped into two categories: major lateral deviations from the route of 50 nautical

Table 3. U.S. ARTCC Codes & Flight Count

ARTCC Code	Location	Sample Flight Count
ZAB	Albuquerque	2014
ZAU	Chicago	3163
ZBW	Boston	1915
ZDC	Wash DC	3348
ZDV	Denver	2157
ZFW	Fort Worth	2570
ZHU	Houston	2617
ZID	Indianapolis	2946
ZJX	Jacksonville	3074
ZKC	Kansas City	2366
ZLA	Los Angeles	2586
ZLC	Salt Lake City	1619
ZMA	Miami	2138
ZME	Memphis	2751
ZMP	Minneapolis	2228
ZNY	New York	2597
ZOA	Oakland	1613
ZOB	Cleveland	3393
ZSE	Seattle	1106
ZTL	Atlanta	3818
Total:		50019

miles and a minor category between 20 and 50 nautical miles. There were 27,300 measurements taken. Of them, 5,264 were determined to be minor deviations with an average lateral deviation from their flight plan of 30 nautical miles; 761 were cataloged as major deviations with an average lateral distance of 73 nautical miles. This translates to about 19% of the total flights having an average lateral deviation of 30 nautical miles and 3% with an average deviation of 73 nautical miles indicating that significant deviations do take place in the European airspace. The study also examined the reason for these deviations and the main cause was attributed to flights cleared to fly direct to a downstream position for fuel and time savings.

For the United States data set, the lateral deviation between route and aircraft position, d_r, was calculated for each ARTCC recording as described above for all the aircraft radar track positions within the associated ARTCC's area of control. The result is a total of 8,111,087 measurements taken from about 50,000 flights. Table 4 summarizes the results by ARTCC. The sample sizes are large but so are the errors. The variability of the data in the form of standard deviation metric ranged from about 10 to 45 nautical miles of lateral deviation. The sample means ranged from about 1 to 7 nautical miles, however the medians (50^{th} percentile) only ranged from -0.01 to 0.08 nautical miles.

Table 4. Lateral Deviation Statistics by ARTCC

Descriptive Summary Statistics						
Airspace Source	Sample Size	Percentiles (nm)			Mean (nm)	Std Dev (nm)
		25^{th}	50^{th}	75^{th}		
United States Airspace: Center Data						
ZAB	427361	-0.399	0.014	0.562	1.352	15.326
ZAU	435974	-0.406	0.050	0.846	2.238	15.706
ZBW	303583	-0.570	0.081	1.700	3.727	22.329
ZDC	565728	-0.319	-0.013	0.356	1.795	13.964
ZDV	490275	-0.267	0.063	0.765	3.770	29.597
ZFW	384097	-0.809	0.039	0.951	1.266	12.797
ZHU	421271	-0.607	0.045	0.890	2.151	20.797
ZID	430507	-0.376	0.059	0.671	1.371	10.867
ZJX	540701	-0.714	0.056	1.100	1.361	11.486
ZKC	443290	-0.571	0.041	0.977	2.943	24.252
ZLA	367723	-0.337	0.014	0.743	5.652	26.665
ZLC	348567	-0.458	0.015	0.545	2.704	22.985
ZMA	377355	-1.000	0.068	2.100	6.397	44.761
ZME	437666	-0.481	0.034	0.844	2.333	18.434
ZMP	404147	-0.417	0.043	0.845	3.548	23.644
ZNY	258725	-0.352	0.057	0.723	1.803	13.986
ZOA	227412	-0.412	0.037	0.726	4.075	22.201
ZOB	472835	-0.418	0.005	0.630	1.356	10.306
ZSE	207031	-0.360	0.038	0.536	2.113	20.574
ZTL	566839	-0.535	0.048	0.820	1.734	13.964
Avg	405554	-0.493	0.039	0.857	2.579	20.816

The difference between the sample median and mean statistics indicates the heavy tailed nature of these distributions. The sample mean is substantially increased by the presence of large lateral deviations on the order of hundreds of nautical miles, while the median is typically unaffected. This observation is not uncommon. In [13], it was independently reported that large tails were present in the lateral measurements collected from flights off the West coast of the United States in Oakland Oceanic and ZLA. In the same paper, a parametric model was successfully fit to the measurements that described two distinct events occurring. Interestingly, it showed some measurements were simply effected by typical deviations from centerline of an aircraft's intended route, while others (large deviations in the tails) were generated by atypical events where aircraft changed route and the automation lacked the updated information.

A reasonable indicator of the magnitude of the typical error behavior described above is the interquartile range (IQR). IQR is the difference between the 75th and 25th percentiles as listed in Table 4. By definition, IQR contains 50% of the distribution. For all 20 ARTCCs, the IQR was on average about 1.4 nautical miles. This is in contrast to the much larger standard deviation which captures both the typical and atypical behavior because it quantifies the entire spread of the distribution. In this case, it captures both the large deviations from the heavy tails and the typical behavior near the center. For this data sample, the standard deviation and IQR are poorly correlated further justifying this claim. Thus, ARTCCs with large standard deviations may not have large IQRs and vice versa.

Cluster analysis is the classification of similar objects into groups [14-15]. For this data set, the 20 ARTCCs are clustered into similar groups in terms of the standard deviation and IQR statistics. This allows us to select a subset of data for detailed analysis to infer claims about a group of ARTCCs. More importantly, it may indicate some common characteristics of clusters associated to the lateral intent process. Lateral intent is one of many input sources for a TP, but it does play a key roll as described previously, and may predict the overall performance of a DST. Graphically, Figure 6 illustrates a two dimensional bubble plot with y-axis in terms of standard deviation and x-axis in terms of IQR statistic. The bubble's size is proportional to the sample size from Table 4 and the three shades denote the three identified clusters.

There are many mathematical approaches in determining the clusters for continuous data [14-16]. Hierarchical clustering divides the data in a successive number of steps where at each step the number of clusters increases until all the data remains in a single group. Utilizing JMP® statistical software package, Ward's minimum variance technique is applied here to produce a number of

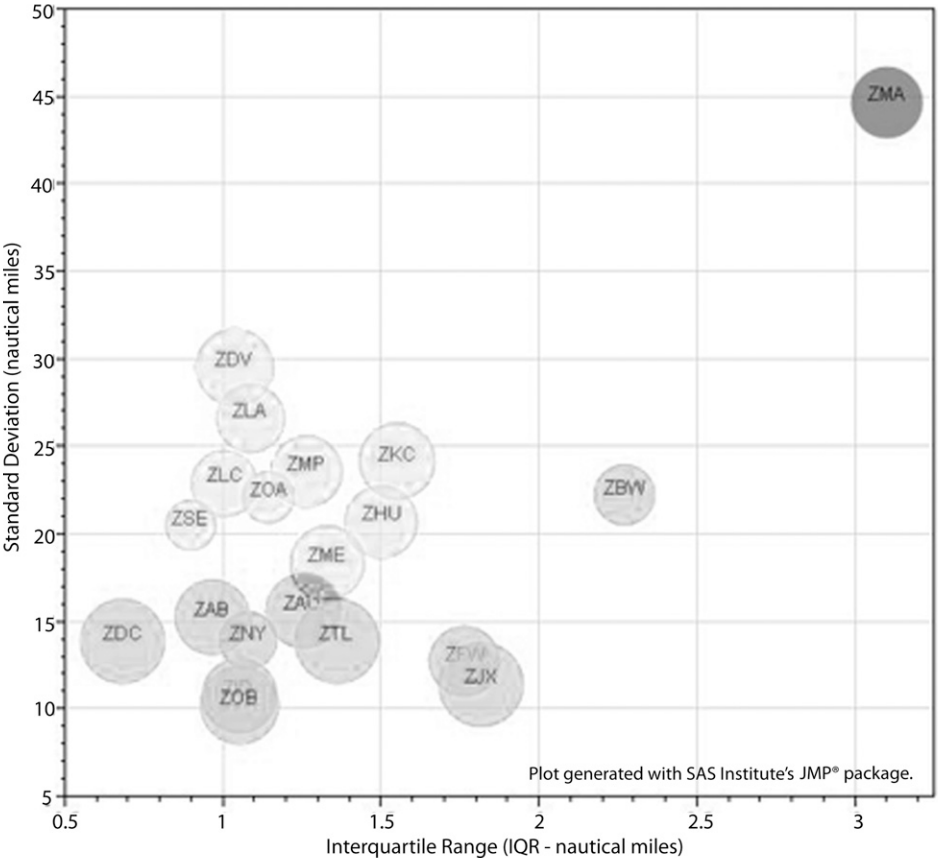

Figure 6. Bubble plot of U.S. ARTCC statistics.

clusters [16]. Figure 7 provides a graphic representation of the results of this technique. This type of figure is called a denogram. It represents a horizontal tree structure with single points as leaves, the final single cluster as the trunk, and the intermediate clusters as branches. The clusters and their grayscale shading in Figure 7 are logically consistent with the illustration in Figure 6.

As shown in Figure 8, the dark gray cluster of ARTCCs (the cluster with the lowest standard deviation) is concentrated in the East coast of the United States. Interestingly, also in the east coast, Miami ARTCC (ZMA) represents the lone cluster with both the largest standard deviation and IQR relative to the other ARTCCs. There could be a number of reasons for this result. Convective or some other severe weather on this particular day could have caused increased reroutes in some areas of the country. The nature of operations and composition of the airspace may play a role in differentiating some facilities or matching them. ZMA in particular may be affected by oceanic traffic from the Atlantic and Caribbean and the limitations of radar coverage over these areas.

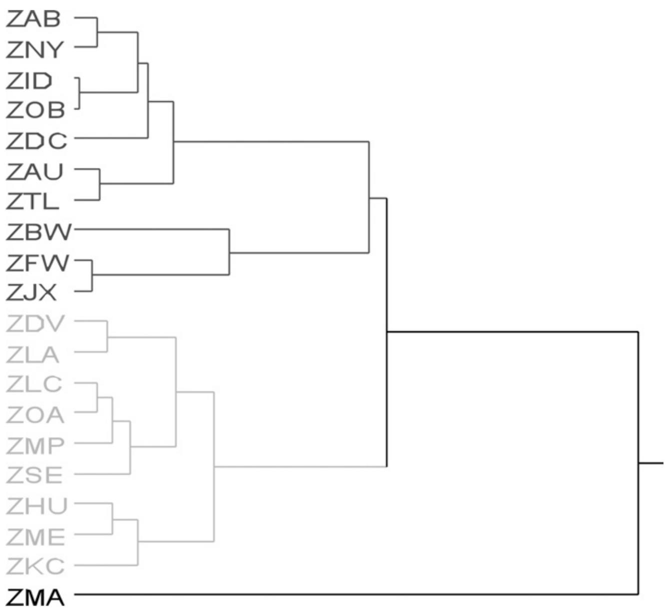

Figure 7. Denogram tree diagram of clusters.

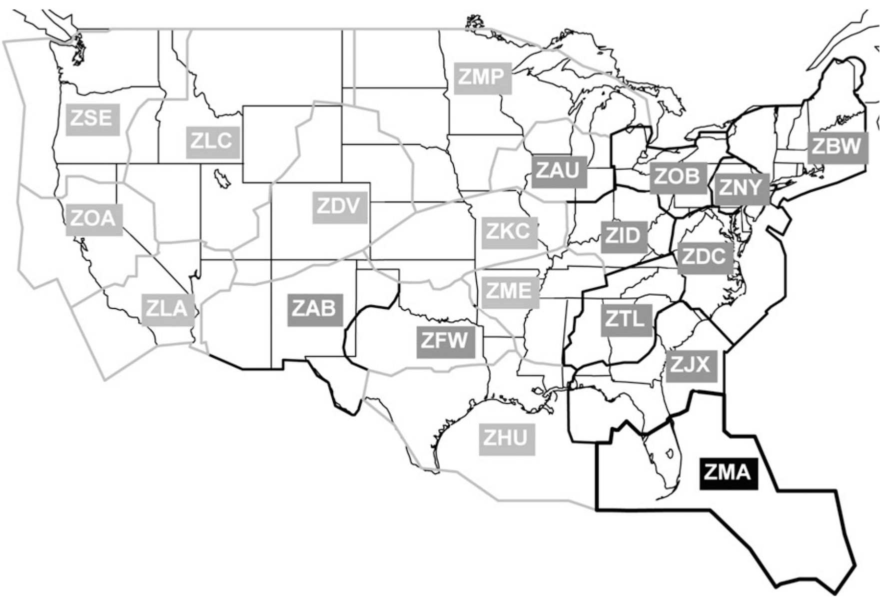

Figure 8. United States map of ARTCC clusters.

Other metrics may give insights into the resulting clusters and distribution of lateral deviations among the ARTCCs. Each flight plan route as described earlier is composed of a series of fix positions and airways. Some of these positions along the route represent turns. More frequent turns in theory may correlate to increased lateral deviations. Aircraft in turns follow more variable curved paths versus straight paths on the rest of the trajectory. Thus, measurements were calculated quantifying the fraction of turn fixes (positions on the flight plan route which formed a turn greater than 30 degrees) compared to the total number of fixes for each ARTCC. The results for our 20 ARTCC sample ranged from 11% to 21%, where the largest was indeed ZMA. However, overall the values did not align with the defined clusters or correlate to a significant degree with the standard deviation or IQR. Between these two statistics, they were slightly more correlated with the IQR values. Since IQR is postulated to be a better estimate of the typical lateral deviation and not a change in route itself, the result is consistent.

It was also postulated that if more amendments were entered, then it would be more likely to capture the lateral intent and lower the resulting errors. To test this, the average number of unique route amendments recorded per flight per ARTCC was calculated. It ranged from approximately 3 to 11 routes per flight with an average of 5 for all ARTCCs. However, the results indicated no correlation to the standard deviation and IQR metrics. Also, there seemed to be no noticeable relationship to the clusters identified. Thus, if amendments have an effect, it is the non-recorded variety that is the suspected cause for the errors being measured in this paper.

Convective weather could play a role on the number of re-routes and thus could influence the lateral intent. Figure 9 illustrates a weather map for the same date of the traffic recording, downloaded from the United States National Oceanic and Atmospheric Administration [17]. It shows the precipitation areas and amounts in North

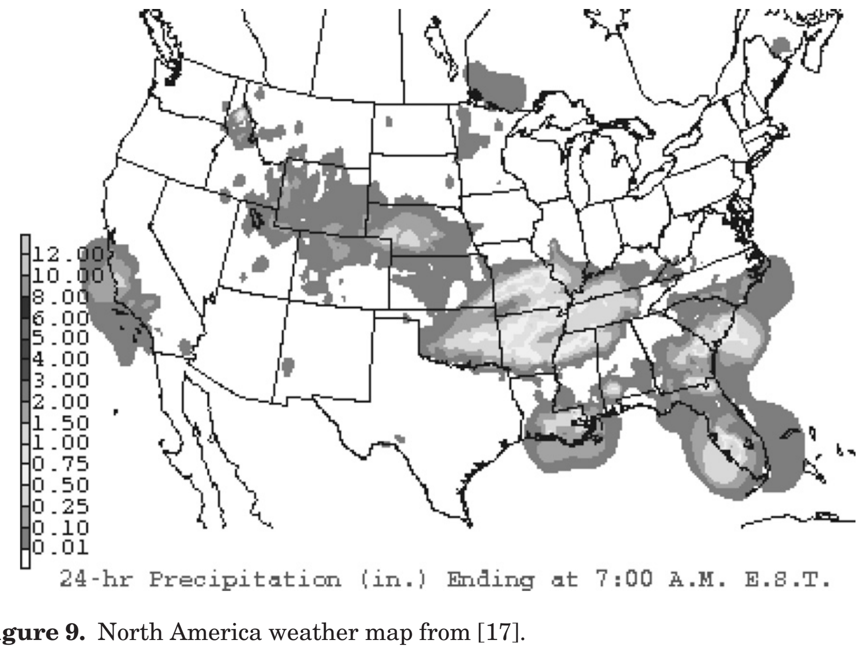

Figure 9. North America weather map from [17].

America during the 24 hours ending at 1200 UTC, with amounts to the nearest hundredth of an inch. From the shading, ZMA did have significant precipitation during the sample date, yet so did other ARTCCs in the southern part of the country and some in the Midwest as well as the west coast.

To get a better understanding within these clusters and the variance between them, ZID, ZMA, and ZMP ARTCCs were selected from each of the three clusters and a detailed comparison of their lateral distributions was performed. This is illustrated in Figure 10. A box plot is depicted that illustrates the spread of the data. The shaded histograms also portray the spread with height proportional to frequency. The inner box represents the 25th, 50th, and 75th percentiles and extending lines referred to as whiskers are 1.5 times the IQR values (length of the box). The separate horizontal lines are the mean values. It is clear from Figure 10 that ZMA has significantly more variability than the other two ARTCCs with ZID having the lowest in terms of total spread, IQR and mean.

Lateral Adherence State Statistics. Lateral adherence states were defined earlier. Figure 11 illustrates the relative frequencies of each adherence state for a subset of ARTCCs. It presents the three selected ARTCC's ZID, ZMP, and ZMA. As shown earlier, the ZMA contains the largest amount of lateral deviation with 58% of the measurements out of conformance overall (i.e. sum of *outerNonConf*

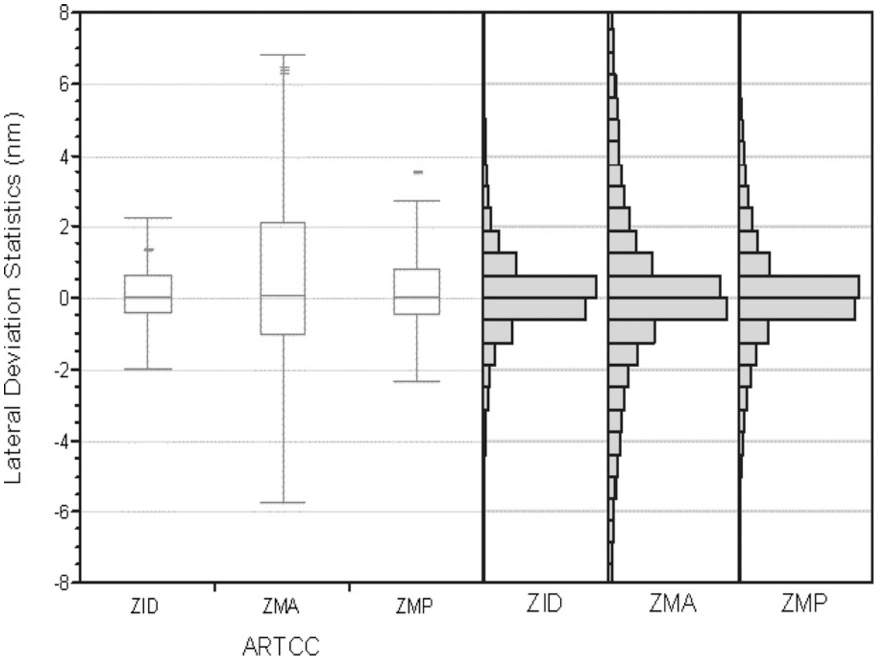

Figure 10. Detailed view on selected ARTCCs.

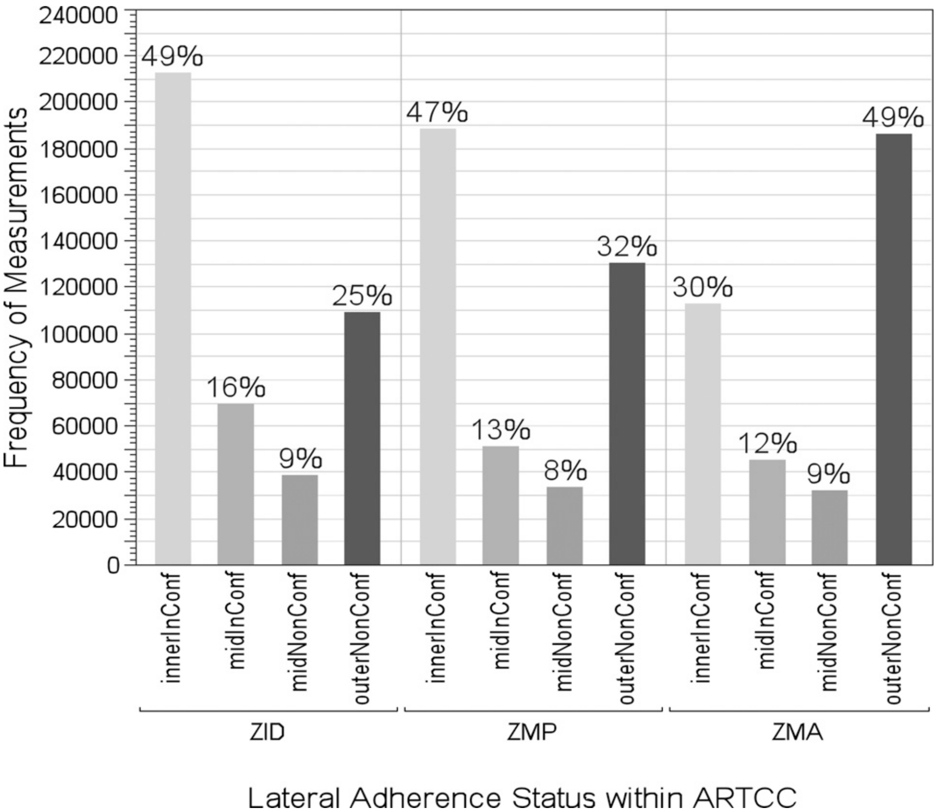

Figure 11. Sample of lateral adherence states.

and *midNonConf* measurements), while ZID has only 34% out of conformance, and ZMP 41%.

These states are intended to provide guidance on the current state of the aircraft in terms of conformance to its current route of flight. The thresholds were already presented in Table 2, they ensure that aircraft labeled in conformance are laterally within one nautical mile of their route. The average IQR reported in Table 4 is reasonably close to the thresholds D_1 to D_3 in Table 2. The lateral adherence results are useful in [11] to determine the type of TP algorithm to apply. In the next section, it is very useful to quantify the impact of lateral deviations on conflict prediction.

Impact of Lateral Error on Conflict Predictions. A conflict occurs when two or more aircraft fly within a defined, typically legally required, separation distance. One of air traffic controllers' core functions is to prevent these events from happening by clearing aircraft to fly trajectories where conflicts cannot occur or amending them to ensure they are resolved. This requires a significant cognitive load involving the controller to maintain a positive mental picture of where the aircraft currently is and where it will fly sometime in the future. The task load is increased by multiple aircraft traveling in different directions both vertically and horizontally. A class of DSTs, called conflict probes (CPs), can aid in this challenging mental process by making automated trajectory predictions, notifying when a conflict may occur in the future, and in more advanced tools offering resolutions. However, these predictions need to be accurate and timely to have utility for the air traffic controller. Furthermore, the JPDO, SESAR, and Australian initiatives all require various conflict probes to perform well for their operational concepts of the future.

Metrics have been defined that quantify the errors associated with these conflict predictions [18-20]. A missed alert error is a conflict between a pair of aircraft not detected at all or not notified within a minimum warning time prior to the conflict's start time, typically five minutes for strategic conflict predictions. A false alert error is a non-conflict event between two aircraft (called an encounter in this paper) that is detected by the CP or represents an alert of a conflict but is removed prior to the conflict occurring. Thus, alerts must be timely and stable to be counted as valid (i.e. correctly detecting an aircraft conflict event with a required lead time).

A CP testing methodology was developed in the late 1990s and documented in [21] that time shifts the recordings of actual aircraft position messages and air traffic control clearances to induce pseudo or test conflict events. The modified traffic recording is then run through the CP as if it were real data. The resulting alerts are matched with the test conflicts generated by the methodology to determine the rates of missed and false alert events from the sample recording. To further illustrate the impact of lateral deviations in this paper, the technique was applied on a flight plan based CP originally developed in [11] and run on a sample scenario taken from Indianapolis ARTCC (ZID).

The sample scenario contains four hours of time-shifted traffic data, amounting to approximately 1100 flights with 139 test conflict pairs for the CP to detect. The overall missed, false, and correct or valid alert quantities are as follows:

- 98 events were valid alerts (VA)
- 41 events were missed alerts (MA)
- 903 events were false alerts (FA)

The overall performance reported is slightly larger than normal because aircraft deviating significantly from the route are normally excluded from the error event counts for strategic or intent based CPs [19-20]. However, this particular study's objective is to quantify the impact of lateral intent errors so excluding them would not allow their measurement.

To evaluate whether lateral deviations influence the CP's accuracy performance, the analysis first examined the two sets of conflict events: those that were correctly predicted (VAs) and those that were missed (MAs). The analysis focused first on the distribution of maximum lateral deviation distance at the start of the conflict for these two sets of data for each of the flight pairs in conflict. However, each event was further partitioned by the reason category in which there are four. A missed alert is an error if no alert was present at all at the actual conflict start time. This was labeled as "NO_CALL_MA". The alternative is the CP did present an alert but within a threshold (five minutes for this study) time of the conflict start time. This is labeled as "LATE_MA". The valid alerts had the remaining two sub-cases. If the alert is presented before the conflict start time but again within the threshold time (same five minutes value as above) yet had a verified reason for being late, its labeled "LATE_VA". These late valid alerts are artifacts of the testing environment, such as a conflict that begins at the start of the traffic sample or within a threshold of a recorded clearance event. These conflicts are considered "pop-up" events and the timeliness requirement is relaxed only for them. The remaining VA events are the standard correct alerts that were correctly matched to a conflict and had the required warning time. These are labeled "STD_VA".

Figure 12 displays the box plot and data points for the maximum lateral deviation for each aircraft pair involved in the four categories of VA and MA events. The late VA events and the no-call MA events had the largest lateral deviations indicating the possible impact lateral deviation has on the CP. Most notably was the contrast of the no-call MA events to the others. The no-call MA had an IQR (range of the box plot) of about 9 nautical miles, while the standard VA was only 2.5 nautical miles.

Like the MA versus VA analysis above, a false alert (FA) analysis was performed as well. The total set of VA events has a lateral deviation mean and standard deviation of 11 and 17 nautical miles contrasted with the FA events of 18 and 29 nautical miles. As illustrated in Figure 13, the IQR (height of the box) is significantly larger for FA events compared to VA events. This provides evidence to support the hypothesis that false alert predictions are induced in part to the lateral deviations of the aircraft the CP is processing.

To further explore the impact that lateral deviations have on the conflict prediction process, a categorical statistical analysis was performed by first generating a 2 by 2 contingency table, illustrated in Table 5. The table partitions conflict events by their lateral adherence state and whether the event was predicted (alert or no alert) by the CP. These alert counts represent the VA and MA events partitioned by their lateral adherence state at the conflict start time. If either flight of the conflict pair was in a state of out of adherence, as defined earlier, then the conflict event was labeled "Out", and "In"

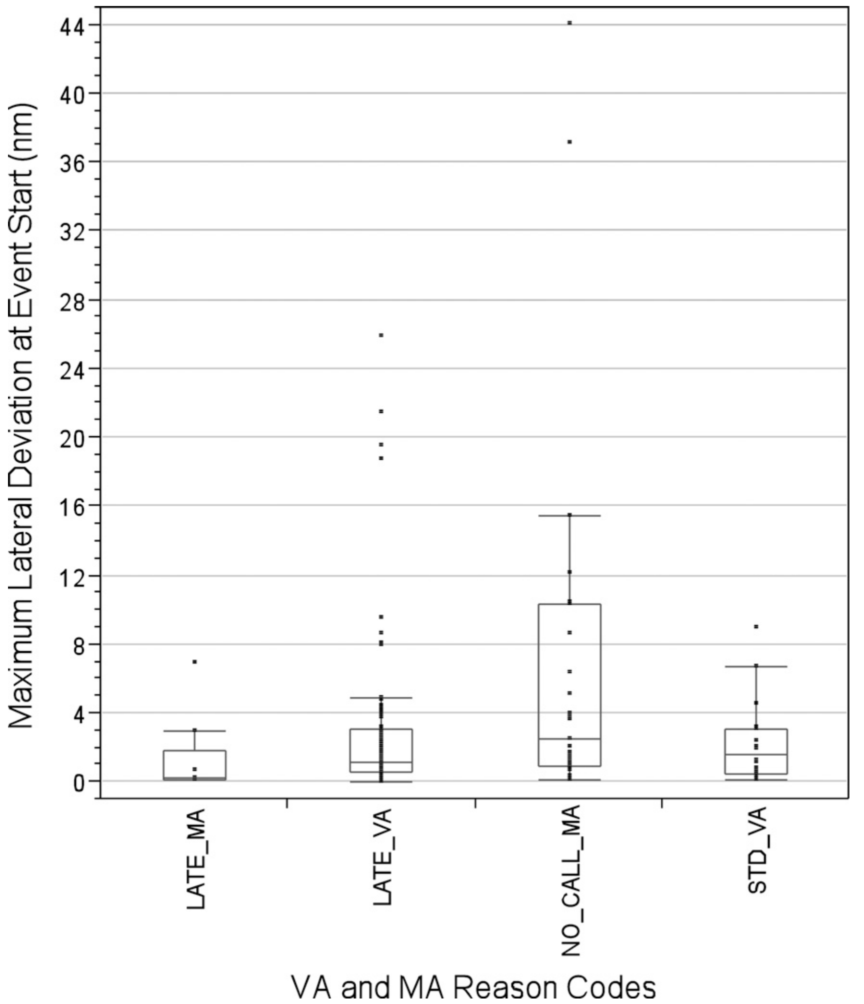

Figure 12. Lateral deviation distribution for VA and MA reasons.

otherwise. Of the total 98 alert events (VAs) 38 were "In" and 60 were "Out". In contrast, the 41 conflicts without alerts (MAs) had 14 that were "In" and 27 were "Out". If the lateral deviations classified by lateral adherence state did not impact the CP's conflict predictions, then the ratio of alerts and non-alerts (VA and MAs) would be the same for both subsets of true conflicts. This can be tested statistically as defined in [22-23] and expressed in equation (1) by calculating the ratio of the squared difference between the expected value of each count and the observed value. If the hypothesis is true, this ratio will follow a chi-squared distribution with one degree of freedom. The expected value is calculated by determining the proportion of total conflict events by the ratio of alert events. For example, the expected VA count is calculated by multiplying the proportion of total conflicts that were in adherence (52 from Table 5) by the total ratio of VA

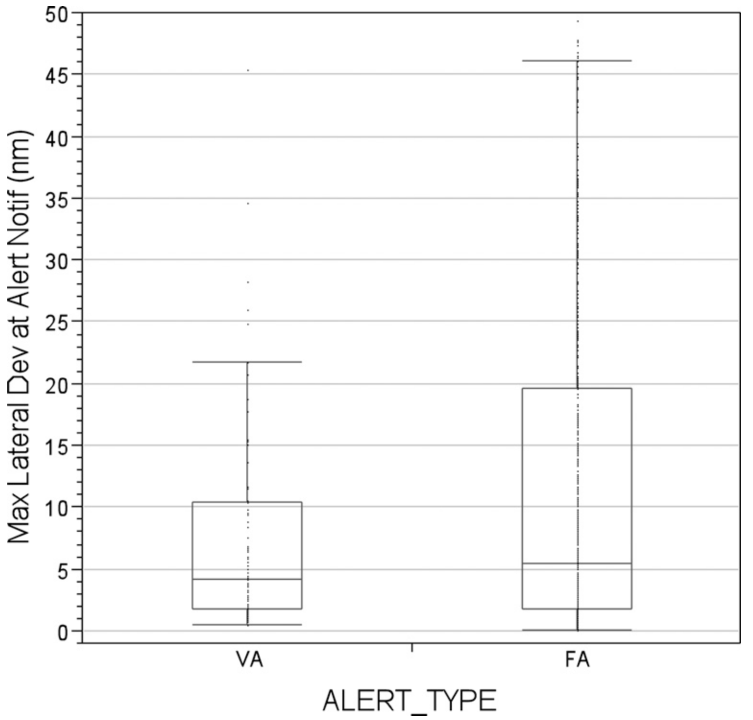

Figure 13. Lateral deviation distribution for VA and FA.

Table 5. Contingency Table for Conflict Events

Alert State	Conflict Event Counts For Each Lateral Adherence State		Totals
	In	Out	
Alert	37 38	61 60	98
No Alert	15 14	26 27	41
Totals	52	87	139
	χ^2=0.265, *df*=1; p-value=0.607		

events (98/139). This results in 37 and listed in the upper right corner cell in Table 5. Thus, application of (1) to all values in Table 5 produces a p-value larger than 0.1, and thus the hypothesis cannot be rejected that the number of MA events are not correlated to an out of adherence state at 0.1 significance level.

The test statistic is χ^2, defined as follows:

$$\chi^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

Where,

O_i is the observed frequency in category i
E_i is the expected frequency in category i

Table 6 illustrates the opposite result for the analogous encounter events (non-conflict) and associated alerts (FAs). The result indicates that more encounter events are alerted than expected for the proportion of encounters that were labeled "Out" and less for "In" adherence. Thus, there is statistical evidence with a p-value near zero to reject the hypothesis that FA events are not affected by lateral adherence.

One additional analysis was performed on the FA events to illustrate an even more direct and overall impact that lateral deviations have on the CP's predictions. A unitless ratio called the min-max ratio was calculated for all non-conflict encounter events and matched to the associated FA events. The min-max ratio is defined in detail in [18]. To summarize, the maximum ratio between horizontal separation and the horizontal separation standard (e.g. 5 nautical miles) and the vertical separation and vertical standard (e.g. 1000 feet) is calculated for each time coincident surveillance position between the aircraft of the encounter. The minimum of all values represents the minimum distance in both dimensions the aircraft pair were separated. Also, if the ratio is less than one, the encounter would be a conflict event. The min-max ratio provides a guide to how close the aircraft came in terms of the separation standards and combines both horizontal and vertical dimensions into one parameter.

For the Figure 14 below, the total number of encounters between each 0.5 min-max ratio starting at 1 was calculated and the associated FA events as well. The fractions of associated FA events to the total encounters per bin were calculated. In Figure 14, the results are plotted with the y-axis as the fraction (estimate of probably of alerts for the bin) and the x-axis is the min-max ratio from 1 to 5.5. The figure's fit curves are power series best fits for these data points.

Table 6. Contingency Table for Encounter Events

Alert State	Encounter Event Counts For Each Lateral Adherence State				Totals
	In		Out		
Alert	175	313	514	376	689
No Alert	2131	1993	2263	2401	4394
Totals	2306		2777		5083
	χ^2=128.22, *df*=1; p-value=0.000				

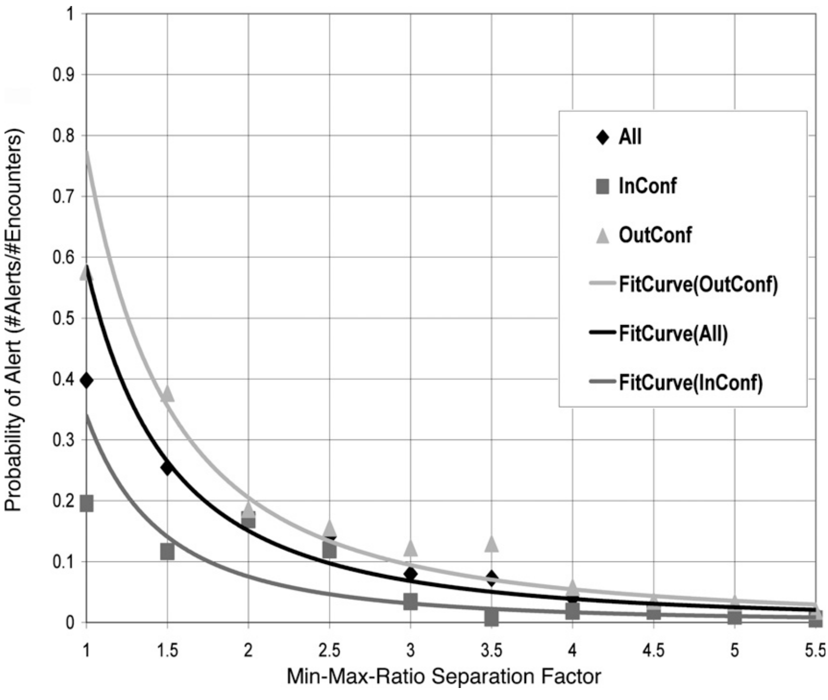

Figure 14. False alert probability curves for lateral adherence state.

The slope of the curve is proportional to the overall sensitivity of the CP to the separation of the encounters it is predicting and the area is roughly equal to the total false alert probability.

Figure 14 fit three curves. The dark gray square labeled curve represents the predictions for alerts that are in lateral adherence at notification time. The light gray triangle labeled curve is for the out of adherence version and the curve with black diamonds represents all the alerts, both in and out of adherence. It is clearly shown that the steepest curve is the in adherence version and least is the out of adherence version with the all curve in between. This gives a good indication that lateral adherence affects the overall sensitivity and thus performance of the CP. It provides direct empirical evidence on the impact of lateral adherence on conflict prediction.

Lateral Deviation Results from Airborne Automation

As envisioned by the JPDO, SESAR, and Australian ATM Strategic planners to some degree, the large lateral errors described in this paper for the ground-based automation systems in both the American and European airspaces have relatively near term solutions. Specifically, an alternate approach to aviation automation both air and ground is to exchange trajectory information constructed and contained within the aircraft's Flight Management System (FMS) computer then utilize this data in the ground system TP. The impact as presented in the CP, described in this paper, is just one example of many of the potential improvement opportunities.

One currently available source of aircraft derived trajectory information is the Automatic Dependent Surveillance Contract (ADS-C). ADS-C is a dependent form of surveillance in which a ground station initiates a contract (dynamic agreement) with an aircraft such that the aircraft will automatically report information obtained from its onboard equipment according to conditions specified in the contract. A Future Air Navigation System (FANS) equipped aircraft can have up to four individual contracts with specific ground stations plus a contract with the Airline Operations Center (AOC).

The current positions in the ADS-C data are based on very precise (relative to ground based radar reports) Global Satellite System (GPS) position reports, and the route positions are exactly what the guidance system within the aircraft's FMS is currently flying to. Like the ground based counterpart presented in the previous sections, lateral deviations between the current ADS-C positions and route positions were calculated from two different sources. This section will report on the lateral deviations supplied by Airservices Australia using data from Australian controlled airspace and similar data in the United States from the FAA's Separation Standards Analysis Team from air traffic collected off the West coast of the United States.

ADS-C's periodic contract specifies the reporting rate and what data groups are requested in each ADS Basic Periodic Report. The following groups can be requested:

- Basic Group containing current position, altitude and time.
- Earth Reference Group containing groundspeed, true track and vertical rate.
- Air Reference Group containing Mach number, true heading and vertical rate.
- Meteorological Group containing aircraft measured wind speed, wind direction and temperature.
- Predicted Route Group (PRG) containing a position and arrival time estimate for the next waypoint and a position estimate for the waypoint that follows.
- Intermediate Projected Intent (IPI) containing position and arrival time estimates for a maximum of ten trajectory change points (not necessarily waypoints, e.g. Top of Descent) ahead of the aircraft.

Besides periodic contracts, there are event contracts and demand contracts. The event contract specifies that for a particular event (e.g. waypoint change event or altitude range deviation event) an ADS report needs to be down linked. The demand contract is a one-time request for an additional Basic Periodic Report.

Australian ADS-C Lateral Deviations. The Australian ADS-C data extracted for this study was collected between February 2008 and January 2009 during the Tailored Arrivals trial performed by Airservices Australia and participating partners. The primary focus of the Tailored Arrivals research is to determine the accuracy and consistency of the aircraft's intended trajectory provided by the Intermediate Projected Intent of the ADS Basic Periodic Report. To eliminate external variables to the maximum extent possible, it was important for the onboard automation (FMS) to fly the aircraft in both lateral and vertical navigation (i.e. LNAV and VNAV) control modes without human intervention.

The ADS-C data was obtained from flights arriving in the early morning, which due to the relative low traffic density were highly unlikely to be subject to air traffic control (ATC) intervention for the arrival. Coupled with the published runway linked Standard Terminal Arrival Routes at the destination, the trajectory of these aircraft can be stable in excess of two hours prior to destination. The consistency of processing these aircraft permits the extraction and analysis of intent data from these flights commencing significantly prior to destination. For consistent results the flight crew were issued with instructions to operate in both LNAV and VNAV modes and ATC were asked not to intervene unless absolutely necessary. The FMS and onboard automation was permitted to operate the aircraft as optimally as possible[1]. Without ATC or pilot intervention the ADS-C position reports of these flights form a consistent and valid data set to analyze lateral deviations from the FMS intended track. The intended or planned track of the aircraft can be constructed from the PRG of the ADS Basic Periodic Report which is consistent with the ground based flight-planned track and actually includes any direct-to clearance programmed into the FMS[2].

To extract the data from these in service aircraft an unmanned duplicate ATC system was established to initiate ADS contracts specifically tailored to the data collection via a separate and additional ADS-C connection. The ADS contract for data collection purposes differed from the ATSP operational contract by an increased reporting rate at two minutes plus supply of all downloadable data. The high reporting frequency was required to analyze the accuracy and consistency of the Intermediate Projected Intent (IPI) over sub-

[1]Operating optimally in this context means operating to a flight-specific Cost Index (CI) determined by the AOC to achieve maximum efficiency in overall network operations.
[2]For purposes of this study it is preferred to construct the reference track from which to determine the lateral deviation from the PRG over the IPI, because of the fixed position of the waypoints in the PRG.

Table 7. Airborne Lateral Deviation Statistics

Airspace Source	Sample Size	Percentiles (nm) 25th	50th	75th	Mean (nm)	Std Dev (nm)
		United States Airspace: ADS-C Data[a]				
U.S.	39012	-0.018	-0.002	0.007	-0.001	0.184
		Australian Airspace: ADS-C Data				
A.A.	26731	-0.019	-0.002	0.015	-0.003	0.026

[a]Adapted from Table 4 in [13]

sequent reports[3]. During the two hour data extraction, at least 60 ADS-C Basic Periodic Reports were received per flight.

All ADS-C data used in this study were obtained from eastbound flights departing from Dubai and Singapore to Melbourne and Adelaide. Data extraction commenced when the aircraft was approximately two hours from destination, typically somewhere around 1000 nautical miles travel distance. The flights were performed by Airbus A330-300, Airbus A340-500, Boeing 747-400 and Boeing 777-300 aircraft (all Honeywell FMS). The data included a total of 778 flights with an average of 34.4 reports per flight. The following lists the break down of flights per aircraft type:

- There were 58 flights of type Airbus A330-300.
- There were 168 flights of type Airbus A340-500.
- There were 258 flights of type Boeing 747-400.
- There were 294 flights of type Boeing 777-300.

As listed in the Table 7 in row labeled A.A., a total of 26,731 ADS-C position reports were analyzed and processed for lateral deviation between the ADS-C Basic Group current position and properly matched the previous PRG next and next plus one route positions. Thus, the lateral deviation is calculated between the aircraft's precise GPS position to the aircraft's matched current FMS known route segment. The results are tremendously accurate compared with the ground-based version reported on in Table 4. The standard deviation and IQR are approximately 800 and 40 times smaller than the U.S. ground-based data results. Translated to feet, the standard deviation amounts to approximately 160 feet and IQR about 200 feet. The histogram depicted the distribution of these errors is illustrated in Figure 15. It forms a fairly symmetric

[3]The position estimates of the trajectory change points in the IPI are given by bearing and distances from the aircraft's current position. The subsequent dynamic conversion to latitude and longitude causes these positions to vary per ADS report. This variation directly influences the lateral deviation as determined with respect to the IPI track.

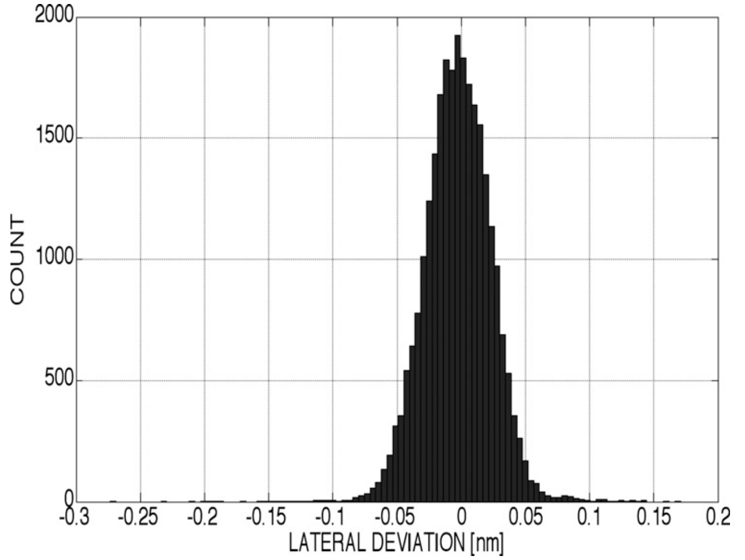

Figure 15. Histogram of lateral deviations of Australian ADS-C data.

distribution about the slightly negative mean and sharply peaked like previous studies [13].

American ADS-C Lateral Deviations. A separate analysis of ADS-C lateral deviations was conducted in 2007 by the FAA's Separation Standards Analysis Team from air traffic collected off the West coast of the United States. The detailed results are published in [13]. The results are summarized in Table 7 within the table row labeled U.S. Like the Australian results, the performance is orders of magnitude improved over the ground based version The sample standard deviation is over 100 times smaller than the ARTCC results, and the IQR is as much as 60 times smaller as well. In units of feet, the standard deviation and IQR translate to about 1100 and 150 feet, respectively. The detailed study in [13] not only reports on descriptive statistics but fits the distribution of errors to a specific parametric model. This model is the beyond the scope of this paper, but indicates that the errors being studied can be mathematically modeled and utilized for simulation experiments of a future ATM system where synchronization of these data sources can be studied further.

Comparing the two sources of ADS-C results indicate some of the differences between the two samples. The Australian data was collected on a sub-set of flights and ATC intervention was purposely excluded by the study or removed in analysis. Additionally, pilots were restricted in the FMS mode of operation they could use for the flights. Thus, large deviations due to changes in the ATC cleared flight plan were typically not present in the Australian data set. For the United States version, the data was collected for 105 days between January and early June of 2007. The data consisted of trans-oceanic flights leaving Oakland oceanic ATC control center and entering the west coast airspace of ZLA (Los Angeles ARTCC). The data was not purposely filtered for ATC deviations or coordinated beyond the normal operational ADS-C process. Thus, the larger difference between IQR and standard deviation in the United States data set and in general larger standard deviation indicates that these outlier events most likely did occur and increased the variance to several times the quantity measured in Australia. However, both ADS-C data sets provide strong evidence of the tremendous improvement in the form of lateral accuracy compared to the ground based versions reported on in the United States and Europe.

CONCLUSIONS

Fostered by the broad next generation ATM initiatives from JPDO, SESAR, and Australia's ATM Strategic Plan, the overall objective of this study was two fold: quantify the lateral deviations between known flight plan routes within the ground and air automation systems across the globe and second determine the impact of these errors on some of the DST functions required for ATC operations. The collaboration of American, European, and Australian researchers provides a broad international perspective and relevance to both the analysis and the source data collected.

Besides the need to collaborate for savings in resources, the particular problem being studied, error of lateral intent in our ATM automation system, is clearly a global issue for all ATSPs. The results from the large United States data collection of 50,000 flights and over eight million measurements reported a standard deviation of approximately 21 nautical miles. The European results cited from [12] reported an average lateral deviation of 30 nautical miles for 19% of the flight measurements taken from a broad data collection of about 27,000 flights across European airspace.

The ground-based results are contrasted with airborne FMS navigated route positions and GPS generated current positions collected using ADS-C. Airborne systems do offer improvements in position accuracy over ground systems, but the dominating factor of improvement is that these systems are more readily updated. This is illustrated from results exhibited in both the United States and Australia, where reported lateral deviation errors are 100 to 800 times smaller than the ground-based version. Australian ADS-C data as listed in Table 7 had standard deviations of approximately 0.03 nautical miles, which translates to less than 200 feet.

ATC personnel as defined in the broad ATM initiatives previously cited will need accurate DSTs to support the complex and safety critical operations they perform. The CP is a DST that directly supports the separation management function by notifying when two or more aircraft are predicted to violate separation standards (i.e. have a conflict) in the future. As described in [24], the uncertainties in aircraft predictions have significant impacts on these ATC functions. To illustrate the lateral error sources in this study, a flight plan or intent based CP was input a set of test scenarios with slightly altered field recordings of actual aircraft flights. About 140 test conflicts and 1100 flights were evaluated to determine if the lateral adherence state had statistically significant impacts on the CP's conflict predictions. For lateral deviations between zero and one nautical mile contrasted against events with larger deviations, the CP's performance had indeed degraded. Other results indicated marginal or inconclusive results for missed alert (not detecting a conflict that really occurs), but the overall sensitivity of the CP's predictions as a function of the separation distance between aircraft was significant and illustrated in Figure 14.

Therefore, the large ground-based deviations reported and impact demonstrated on a CP tool show lateral deviations are a key source of error in our ground-based TP process, core to many of our DST functions. Synchronization with airborne data sources like ADS-C offers a reliable and tremendously accurate solution to improving aircraft predictions in air traffic control.

Overall, the international collaboration that took place to perform this study is the type of global cooperation that will be needed to address the challenging ATM problems faced by all nations and ATSPs. The study reports on one aspect of the TP process highlighted earlier. Vertical deviations, time based errors, and weather forecasts mark only a few that can continue to be studied in the same manner set forth in this paper.

ACRONYMS

AA	Airservices Australia
ADS-C	Automatic Dependent Surveillance–Contract
ANSP	Air Navigation Service Providers
AOC	Airline Operations Center
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATM	Air Traffic Management
ATSP	Air Traffic Service Providers
CNS	Communications, Navigation, and Surveillance
CP	Conflict Probe
DST	Decision Support Tools
EFICAT	EUROCONTROL Flight Information Consistency Analysis Tool

FA	False Alert
FANS	Future Air Navigation System
FMS	Flight Management System
GPS	Global Satellite System
HADDS	Host Air Traffic Management Data Distribution System
IPI	Intermediate Projected Intent
IQR	Interquartile Range
JFK	John F. Kennedy International Airport
JPDO	Joint Development Planning Office
MA	Missed Alert
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
Nm	Nautical miles
PRG	Predicted Route Group
SESAR	Single European Sky ATM Research Initiative
TP	Trajectory Predictor
US	United States
UTC	Coordinated Universal Time
VA	Valid Alert
ZDC	Washington ARTCC
ZID	Indianapolis ARTCC
ZLA	Los Angeles ARTCC
ZMA	Miami ARTCC

REFERENCES

[1] Joint 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.
[2] EUROCONTROL, "Single European Sky ATM Research (SESAR) In Brief, Delivering the Future ATM System in Partnership," [online library], URL: http://www.eurocontrol.int/sesar/, October 2008.
[3] S. Mondoloni, M. Paglione, S. Swierstra, C. Garcia-Avello, S. M. Green, "A Structured Approach for Validation and Verification of Aircraft Trajectory Predictors," 23rd Digital Avionics Systems Conference, Salt Lake City, Utah, October 2004.
[4] S. Mondoloni, S. Swierstra, "Commonality in Disparate Trajectory Predictors for Air Traffic Management Applications," 24th Digital Avionics Systems Conference, Washington D.C., October 2005.
[5] S. M. Green, , R. A. Vivona, M. P. Grace, T. C. Fang, "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, Boston, MA.,1998.
[6] D. Brudnicki, W. Arthur, K. Lindsay, "URET Scenario-based Functional Performance Requirements Document," MTR98W0000044, MITRE CAASD, April 1998.
[7] S. Mondoloni, "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.
[8] K. Lindsay, "Results of a URET Operational Utility Experiment," MTR99W0000081, MTRE CAASD, January 2000.
[9] M. Paglione, R. D. Oaks, H. F. Ryan, S. S. Summerill, "Description of Accuracy Scenarios for the Acceptance Testing of the User Request Evaluation Tool (URET) /

Core Capability Limited Deployment (CCLD)," Federal Aviation Administration, William J. Hughes Technical Center, Atlantic City, NJ, January 27, 2000.
[10] R. D. Oaks, M. Paglione, "Determination of Lateral Flight Adherence in Recorded Air Traffic Data," American Institute of Aeronautics and Astronautics Guidance, Navigation, and Control Conference, San Francisco, CA, August 15-18, 2005.
[11] M. M. Paglione, C. Santiago, A. Crowell, R. D. Oaks, "Analysis of the Aircraft to Aircraft Conflict Properties in the National Airspace System," American Institute of Aeronautics and Astronautics Guidance, Navigation, and Control Conference, AIAA 2008-7143, Honolulu, Hawaii, August 18-21, 2008.
[12] EUROCONTROL, "Flight Data Metrics, Overall Report on the 15 November 2006 Data Collection," CGP07/86/35.08, Contract: EFICAT, July, 2007.
[13] C. Gerhardt-Falk, L. Martin, S. Ellis, "Correlation of Airborne Position Estimates to Ground Based Independent Estimates and Deviations from Flight-Planned Tracks," American Institute of Aeronautics and Astronautics Guidance, Navigation, and Control Conference, AIAA 2007-6520, Hilton Head, South Carolina, August 2007..
[14] B. S. Everitt, S. Landau, M. Leese, Cluster Analysis, 4th ed., Oxford Press 2001.
[15] L. Kaufman, P. J. Rousseeuw, Finding Groups in Data, An Introduction to Cluster Analysis, John Wiley and Sons, 2005.
[16] SAS Institute Inc., JMP Statistics and Graphics Guide, Release 7, 2007.
[17] United States National Oceanic and Atmospheric Administration, NOAA Central Library U.S. Daily Weather Maps Project, [online library], URL: http://www.hpc.ncep.noaa.gov/dailywxmap/index.html.
[18] M. Paglione, M. Cale, H. Ryan, "Generic Metrics for the Estimation of the Conflict Prediction Accuracy of Aircraft to Aircraft Conflicts by a Strategic Conflict Probe Tool," Air Traffic Control Quarterly, Vol. 7 (3), Fall 1999.
[19] K. Bilimoria, "A Methodology for the Performance Evaluation of a Conflict Probe," Journal of Guidance, Control, and Dynamics, Vol. 24 (3), May-June 2001.
[20] M. Paglione, R. Oaks, H. Ryan, "Methodology for Evaluation and Regression Testing a Conflict Probe," Digital Avionics System Conference, October 2004.
[21] M. Paglione, R. D. Oaks, K. D. Bilimoria, "Methodology for Generating Conflict Scenarios by Time Shifting Recorded Traffic Data," American Institute of Aeronautics and Astronautics Technology, Integration, and Operations (ATIO) Technical Forum, November 2003.
[22] Kachigan, Statistical Analysis, An Interdisciplinary Introduction to Univariate and Multivariate Methods, Radius Press, 1986.
[23] A. Agresti, Categorical Data Analysis, 2nd ed., John Wiley and Sons, 2002.
[24] D. McNally, D. Thipphavong, "Automation Separation Assurance in the Presence of Uncertainty," 26th International Congress of the Aeronautical Sciences, 2008.

BIOGRAPHIES

Mike M. Paglione graduated from Rutgers University College of Engineering, in New Brunswick New Jersey in the U.S.A. with a B.S. degree in Industrial Engineering in 1991. He obtained a M.S. degree in Industrial and Systems Engineering in 1996 from Rutgers University Graduate School., in the U.S.A.

He is the Conflict Probe Assessment Team Lead in the Federal Aviation Administration's Simulation and Analysis Group at the FAA W. J. Hughes Technical Center, Atlantic City, New Jersey. He was FAA's Rutgers University Fellow from 1994-1996, Accuracy Test Lead for the FAA's User Request Evaluation Tool, Program Manager for the Joint University Program from 1999 to 2004, currently project lead on the Automation Metrics Test Working Group (a cross organizational team developing and implementing metrics for the En Route Automation Modernization Program), and a local team lead supporting a NextGen project investigating improvement to the separation management functions in the en route automation.

Mr. Paglione is a senior member of the American Institute of Aeronautics and Astronautics (AIAA) and has numerous publications in the area of air traffic management automation, especially in the evaluation and testing of operational systems.

Ibrahim Bayraktutar graduated from Middle East Technical University, Turkey with a B.S. degree in Industrial Engineering in 1990. He obtained a M.S. degree in Systems Analysis in 1992 and a M.S. degree in Operations Research in 1993, both from Miami University, U.S.A.

He has been working for EUROCONTROL since 1994 participating in various areas regarding Air Traffic Management, specializing in Trajectory Prediction and Management. He is currently managing Trajectory Management Framework activities in ATC Operations and Systems unit. He has several publications on Systems Analysis, Systems Dynamic Modelling and Trajectory Prediction/Management (e.g. Impact of Factors, Conditions and Metrics on Trajectory Prediction Accuracy with Stephane Mondoloni, CSSI Inc. at FAA/EUROCONTROL R&D Seminar, Baltimore, U.S.A. in 2005).

Mr. Bayraktutar is the co-chair of EUROCONTROL-FAA Action Plan 16 on Common Trajectory Prediction Capability. He is also a member of ORSA (Operations Research Society of America).

Greg McDonald commenced work as an Australian Air Traffic Controller in 1981. He obtained an associate diploma of applied science in Computing in 1993 and a degree in Computing in 1996, both from Monash University, Australia.

He has been working for Airservices Australia and its predecessors since 1981 in various roles including radar, non radar sectors plus tower and Search & Rescue. Since 1995 he has worked in an Operational Support role participating in various activities necessary to ensure Australian ATC is conducted efficiently and safely among which was the development of the Australian ATM Strategic Plan. He was the project manager for AUSOTS the Australian Flex Track initiative designed to deliver route efficiencies to airlines operating in the Australian Environment. He is currently managing the Tailored Arrivals Trial in Australia and examining the accuracy and possible ground system use of aircraft derived trajectory data.

Mr. McDonald is Australian representative for EUROCONTROL-FAA Action Plan 16 on Common Trajectory Prediction Capability.

Jesper Bronsvoort is an Aerospace Engineer currently working with Airservices Australia as part of his postgraduate education in Aerospace Engineering at Delft University of Technology, the Netherlands. He holds a cum laude BSc. degree in Aerospace Engineering from Delft University of Technology, The Netherlands (2006).

In his postgraduate, he particularly focuses on the field of optimization of flight operations. He is currently performing research into the use of aircraft derived data to improve trajectory prediction for decision support tools.