Objective Assessment Method for RNAV STAR Adherence

Michael Stewart
San José State University Research Foundation
NASA Ames Research Center
Moffett Field USA
Michael.J.Stewart@nasa.gov

Bryan Matthews
Stinger Ghaffarian Technologies Inc.
NASA Ames Research Center
Moffett Field USA
Bryan.L.Matthews@nasa.gov

Abstract— **Flight crews and air traffic controllers have reported many safety concerns regarding area navigation standard terminal arrival routes (RNAV STARs). Specifically, optimized profile descents (OPDs). However, our information sources to quantify these issues are limited to subjective reporting and time-consuming case-by-case investigations. This work is a preliminary study into the objective performance of instrument procedures and provides a framework to track procedural concepts and assess design specifications. We created a tool and analysis methods for gauging aircraft adherence as it relates to RNAV STARs. This information is vital for comprehensive understanding of how our air traffic behaves. In this study, we mined the performance of 24 major US airports over the preceding three years. Overlaying 4D radar track data onto RNAV STAR routes provided a comparison between aircraft flight paths and the waypoint positions and altitude restrictions. NASA Ames Supercomputing resources were utilized to perform the data mining and processing. We assessed STARs by lateral transition path (*full-lateral*), vertical restrictions (*full-lateral/full-vertical*), and skipped waypoints (skips). In addition, we graphed frequencies of aircraft altitudes relative to the altitude restrictions. *Full-lateral* adherence was always greater than *Full-lateral/full- vertical*, as it is a subset, but the difference between the rates was not consistent. *Full-lateral/full-vertical* adherence medians of the 2016 procedures ranged from 0% in KDEN (Denver) to 21% in KMEM (Memphis). Waypoint skips ranged from 0% to nearly 100% for specific waypoints. Altitudes restrictions were sometimes missed by systematic amounts in 1,000 ft. increments from the restriction, creating multi-modal distributions. Other times, altitude misses looked to be more normally distributed around the restriction. This tool may aid in providing acceptability metrics as well as risk assessment information.**

Keywords-data mining; RNAV STAR; procedures; adherence; waypoint

I. INTRODUCTION

Area navigation (RNAV) is a cornerstone of the Federal Aviation Administration's (FAA) plan for future instrument procedures. Between 2009 and 2016, 264 RNAV standard terminal arrival routes (STARs) were implemented, and they are still continuously increasing [1]. This rapid implementation has incurred unforeseen side effects. Major and regional airlines have both voiced safety concerns about deviations at a NASA workshop meeting in November 2015, Boston's STARs were redesigned in 2016 after numerous operational problems, and Atlanta was forced to stop using the vertical and speed profiles of their newly designed RNAV STARs [2],[3]. Our interest in RNAV procedures is pointed specifically towards RNAV OPDs and stems from their increased functionality (e.g., vertical profiles and speed control) and its resulting complexity. To fully understand the operational effects of this paradigm change, either from the flight deck or air traffic control perspective, we need a source of objective data describing operational performance of RNAV procedures. That is, a method for comprehensibly capturing and describing the adherence trends of instrument procedures. We introduce the idea that procedures are important entities in their own right and should be treated as the unit of analysis. They are not just collections of waypoints, and it is performance along a path through a STAR that should be analyzed. By monitoring the flightpaths of aircraft, we can understand how STARs are functioning in the airspace. Currently no system exists that can monitor the adherence of aircraft utilizing RNAV procedures in a comprehensive way.

RNAV STARs are often designed to enhance efficiency and regulate throughput. This expands upon the previous intentions of instrument procedures: terrain avoidance and standardized routing. This expansion of functionality increases complexity, particularly, the addition of the speed restricted vertical profile for the precision it requires, and the addition of transition routes to control specific runway assignments. Vertical profiles add variables pertaining to aircraft performance such as drag, idle, and auto-flight variability; transitions add rigidity by limiting lateral path flexibility.

A. Purpose

In order to maintain or reduce the current level of risk we need to understand the implications of existing and new procedures and our system of controlling air traffic. A possible risk with deviating from a STAR is loss of separation leading to a midair collision. Complying with a STAR keeps the airplane on a known trajectory so it can be protected. Issues such as energy management, loss of control, controlled flight into terrain, wake turbulence, and weather are also possible areas of concern for RNAV STAR adherence. We do not

understand all of the reasons that lead to non-adherence and this deficit in knowledge can increase risk.

This research is a preliminary overview of adherence trends and a demonstration of our data mining methodology to further the knowledge of the air traffic system. By looking at Instrument procedures as entities, we can investigate the design characteristics they embody and observe how our current air traffic behaves when guided by those concepts. In addition, we can continually monitor our airspace system for risk factors to detect changes as they occur.

B. Background

First, we needed to understand how to describe the adherence of a STAR and which design characteristics and external variables might impact flying a procedure. We used the Aviation Safety Reporting System (ASRS) to identify error categories, so we could see which variables in the procedures, environment, and human operators contribute to operational irregularities. We are investigating objective performance, and to do this we needed to understand how human error could be associated with procedure design and outside influences. Users of RNAV procedures are limited to two groups: pilots and air traffic controllers. These groups differ in goals and tasks, so we needed to investigate them separately. Therefore, we reviewed ASRS reports of RNAV procedures from each group's perspective in order to understand the differences. In addition to ASRS we reviewed literature on RNAV arrival concepts. Much of the previous research was centered on air traffic [5],[6].

1) Pilot ASRS Review: We coded 377 ASRS pilot reports related to RNAV STARS. We read these reports, then the issue or deviation and its reported cause was recorded. Issues mostly fell into three categories: altitude, track, and speed deviations. Overall, reporting of deviations is known to under-represent actual occurrence rates and should be considered here. Speed deviations were the least reported issue, most likely because they are difficult to detect by ATC. Some of the categories were not the sole cause of a problem; therefore, some categories were counted more than once for a given report.

2) Vertical Deviations: Vertical deviations happen when an aircraft is not at the altitude expected by ATC. 35% of vertical deviations were associated with autopilot and flight management system (FMS) confusion. ATC communication, vectoring, and STAR modification accounted for 26%. Procedure design accounted for 13%, and charting misreads accounted for 9%. The remaining 17% were caused by manual pilot errors and other unknown reasons.

One of the most commonly reported ATC interventions was runway reassignment (i.e., changing the procedure during its use). This requires the flight crews to reprogram the FMS and change autopilot modes to maintain the correct flight path. Time pressure was often reported as a factor in making FMS errors. When procedure design was reported as a cause, steepness of the vertical profile, number of restrictions, and deceleration requirements were noted as contributory. In addition, tailwinds were mentioned for impeding the ability to maintain the fixed vertical path with the increased ground

speed. Charting misreads issues included notes, clutter, and electronic interface problems.

3) Lateral Deviations: Lateral or track deviations occur when the airplane departed its expected lateral path and had three main causes: ATC intervention, weather avoidance, and FMS errors. However, similar to the vertical deviations, the groups were not mutually exclusive, and there were many reports that initially started with a controller amending a previously cleared flight path, followed by the crew making an error with the FMS.

C. Literature Review

In addition to our ASRS review, we began our literature search with another ASRS investigation. Barhydt and Adams [2006] reviewed 124 reports to categorize and count incidents. They found that four categories accounted for most of the reports. These were: ATC procedures, airline operations, aircraft systems, and procedure design [4].

In addition to ASRS reports, Barhydt and Adams [2006] used subject matter experts (SMEs) to uncover RNAV procedure usability issues. These pilots and controllers explained issues they felt contributed to problems. Issues cited were: ATC procedures, flight crew procedures, training, clearance discrepancies, aircraft instrumentation, FMS databases, and procedure design standards [4]. Many of these issues were noted as branches of a high-level issue: workload from procedural complexity.

The constructs of complexity and workload are variable between individuals, but performance degradation and flight path deviations are sometimes quantifiable. Smith [2005] investigated the construct of controller workload. Controller human-in-the-loop simulations were used to compare traditional ground-based procedures with RNAV STARs containing vertical profiles. Dependent variables for the construct of workload were: number and content of radio communications. Rationale for reducing radio communication was to reduce misinterpretation, which can add risk. Therefore, the reduction of this information transfer method would be beneficial [5]. The results of the study showed a significant reduction in communications and their content. However, the author noted that workload might not be reduced due to monitoring requirements still required by the controller. In addition, the reliability of verbal communications compared to using a chart to deliver information has not been investigated to our knowledge. For risk reduction, subsequent errors rates between conditions could be a useful metric. Since the study was a simulation of Atlanta airspace, there may be differences in other cities. Additionally, weather and pilot deviations could increase the level of communications during an RNAV procedure and should be understood. Lastly, the author noted that passive controller monitoring roles are increased with RNAV procedures and could be detrimental to controller performance.

Metzger and Parasuraman [2001] investigated how controllers were able to detect conflicts while actively controlling, versus passively monitoring traffic. Since a procedure is a preplanned set of instructions, they result in the controller being a passive monitor of aircraft flying those

instructions. RNAV procedures can be in excess of 170 miles and may contain numerous "instructions" in the form of altitude and speed restrictions. Therefore, the opportunity for non-adherence is increased and the controller must monitor more parts of the procedure for a longer period of time. During high traffic conditions, passive monitoring increased the time it took controllers to identify conflicts [6]. If this phenomenon is consistent it could result in reactive controller responses versus planned and coordinated separation.

Olson and Olszta [2010] [7] discuss safety in regards to TCAS events measured in the terminal airspace. The study finds that procedures that are designed to level-off within 1,000 ft between departures and arrival paths would trigger TCAS RA events. In this scenario, the TCAS events were occurring in protected space, however, if a flight is not adhering to the vertical restrictions of the procedure these events would become more serious. From the ASRS review vertical non-adherence was reported to be an issue, and when coupled with TCAS events becomes a central motivation for assessing the vertical adherence.

In all of these discussed scenarios it is impossible to determine the exact cause without reviewing pilot/controller voice communications. However, these events highlight the need to investigate the performance of the RNAV STARs and quantify the occurrence rates where the risk of these events are raised.

II.   METHOD

A. Materials

The architecture to capture these adherence characteristics was built in Python using the Apache Spark large-scale data processing framework [1] and deployed at the NASA Advanced Supercomputing facility[2]. The distributed jobs utilized 65 nodes, each with 2 x 8 core 2.6 GHz processors and 32GB of memory. The jobs took approximately 30 mins per airport up to 2-3 hours for the higher traffic airports.

Python was used for statistical analysis and plotting. Tableau Desktop data visualization software was utilized for making different graphs and renderings.

B. Data Sources

1) *Aircraft Flight-track Data:* The primary data source utilizes recordings of radar tracks from the Center TRACON Automation System (CTAS). CTAS is a system used by the FAA's Air Traffic Control Center (ARTCC) facilities and is designed to provide controllers with automated tools for planning and controlling arrival air traffic [8]. A byproduct of the system is the recordings of 4D radar trajectories throughout the NAS. For this study 17 of the 20 ARTCC facilities were available over three years (2014, 2015, and 2016). The three ARTCC facilities not covered were ZID, ZJX, and ZMA. Stored data includes time, latitude, longitude, altitude, and ground speed and were approximately 1.8TB in size with over 14 million arrivals landing at 24 airports.

2) *Environmental Data:* The NOAA Rapid Refresh (RR) data[3] were parsed over the same three-year period to capture snapshots of winds and temperatures aloft. A subset of these measurements was gathered at 1,000 ft. intervals ranging from 4,000 - 40,000 ft. above each destination airport at a one hour sampling rate. This data source was over 2.7TB.

3) *Procedure Data:* STAR characteristic data are sourced from the Coded Instrument Flight Procedures (CIFP). These data include RNAV STAR waypoints (latitude/longitude), airspeed restrictions, altitude restrictions, and descent gradient (degrees) for each waypoint. Several other derived variables are included for further investigative capability in TABLE I. This data source was approximately 1.5GB.

C. Data Processing

All STAR characteristics were extracted from the CIFP data and concatenated into unique transition paths for each input and output transition combination (i.e., if there were three input transitions and two runway transitions the total unique paths created would be 6 for that STAR). See Figure 1 for an illustration. The altitude and speed restrictions were extracted and linked to each waypoint in the route and stored for each transition path. This process was done for each (56 day) chart cycle to track changes yielding a comprehensive list of unique transition paths across the STARs for all intervals. These paths were used to determine whether a flight was following the route's published lateral track. To accomplish this, each flight's trajectory in the CTAS data was compared against each of the route's unique paths. Lateral adherence was defined to be a flight that crossed inside a 1.0 nm radius around each waypoint. This is in accordance with the RNAV1 requirements of the procedures. All other flights that did not have full waypoint adherence throughout a transition path were considered not fully laterally compliant. Flights that were laterally compliant with at least two waypoints on the transition path and compliant with the first and last waypoint in the path were identified as having skipped waypoints and the skipping behavior was analyzed. All other non-laterally compliant flights were excluded from this analysis due to the ambiguity in the entry and exit points.

Due to the slower 12-second radar sweep, a flight could fly inside the 1.0 nm radius at a waypoint while having its radar hit outside of the threshold. To resolve this issue, flight paths were up-sampled by linear interpolation to every five seconds between samples. Flights were checked against longer transition paths first to avoid matching the flight to a shorter

TABLE I List of variables computed or extracted from the data.

| Aircraft	Procedure	Environment	Airport
• Unix Time/Date • Aircraft Type • Equipment Suffix • Ground Speed (Kts) • Altitude (ft) • Decent Rate (ft/min) • Slope to next waypoint (degrees) • Above/Below Vertical Excursion (ft) • Flight Plan • Percent Route Complete (%)	• Waypoint Name • Altitude Restrictions (ft) • Waypoint Type (At, Window, Above, Below) • STAR Name (iteration) • High/Low Slopes between waypoints (degree) • Transition Counts	• Tailwind Component (Kts) • Temperature (C)	• Arrival Rates (last 15 mins/next 15 mins)
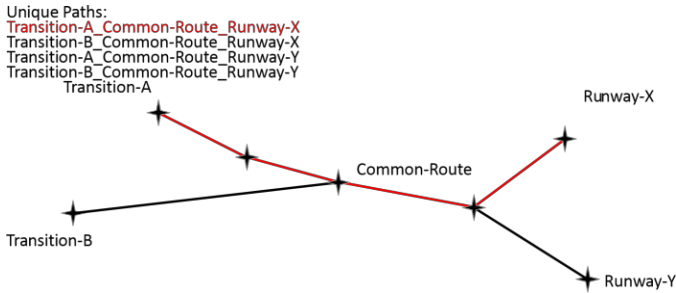
Figure 1. Illustrates the unique transition paths from a STAR.

transition path that may be a subset of a longer one. Once a path was found to match with full lateral adherence, a snapshot of the waypoint and flight's features were logged at each waypoint crossing. Altitude excursions above or below the RNAV defined restrictions for the waypoint ±300 ft. were logged. Anything within the published altitude restrictions were considered vertically compliant and marked to have a 0 excursion. The winds aloft at the altitude of the aircraft at the waypoint crossing were retrieved from the parsed RR data. Using the wind vector and the heading of the flight traveling along the published route, the tail wind component was computed and logged at each waypoint crossing. Additional external factors such as arrival rate at the airport over the last 15 mins were computed from the CTAS data and logged at each waypoint crossing. See TABLE I for a more comprehensive list of the parameters captured from the data

D. Design

This study was based on archival data mining. The originality of the work was in the structured combination of data sources included: radar tracks (aircraft), procedure characteristics (procedure), weather (environment), and airport arrival rates (airport). These data sources were parsed and merged together to create a comprehensive set that can be used to show trends and insights in the adherence of these procedures.

4) Aggregate Adherence Rates: We used the median to describe the central tendency of the adherence distributions due to skew; variability of the data was described using interquartile range (IQR) of Q1 to Q3 (i.e., 25% either side of the median). Yearly aggregate data comprise 365 sample proportions rounded to the nearest one percent. Daily samples are unique flights of interest for all STARs at the airport divided by the total arrival count. Time series data use the same method as previously described, but is smoothed using a 7 day moving average filter.

5) Skipped Waypoints: Waypoint skip proportions were calculated by dividing the number of flights that were counted as skipping the waypoint by the total number of flights that would have passed by or crossed over the waypoint on a transition path. Single waypoints are sometimes part of more than one transition. Therefore, they may be listed more than once with different proportions. A flight may skip multiple waypoints within a transition path, however it must cross at

least two waypoints in a row and also cross the first and last waypoint in the transition path to be included in the tallies. The skipped waypoint rates were computed for each waypoint. Waypoints on a transition that did not have at least 1,000 flight crossings, laterally compliant or not, were ignored due to low sample size.

6) Vertical Characteristics: We looked at vertical adherence trends only if the flight completed 100% of a lateral path for a given transition. It was assumed that the vertical path would be advantageous if the flight was laterally compliant. A 300 ft. error margin was given above and below a restriction altitude and flights were counted as "compliant" if within the margin. In addition, we used proportions rather than counts to normalize for differences in waypoint traffic volume.

We investigated types of waypoint restrictions and categorized them (e.g., at or below) to compare aggregate adherence proportions. In addition, we compared the altitudes aircraft were crossing the waypoints and created histograms to visualize the data. Altitudes were binned in 100 ft. increments from the waypoint restriction altitude.

III.   EXAMPLE DATA

A.   Full Procedure Adherence Rates

Overall median adherence of the *full-lateral* and *full-vertical* profiles (i.e., 100% adherence) for 24 of the most heavily-trafficked airports in the data set (see Figure 2). The median adherence ranges from 0% in KDEN (Denver) to 24% in KMEM (Memphis). In addition, a time series graph of these two airports shows the variability of the data as well as the differential between *full-lateral* only and *full-lateral/full-vertical* adherence. These trends can be seen in Figure 3 and Figure 4. The vertical tick marks represent the 56 chart cycles from 2014-01-01 through 2016-12-31. The rates of adherence still vary considerably over time. Since *Full-lateral/full-vertical* is a subset it is less frequent than *full-lateral- only*.
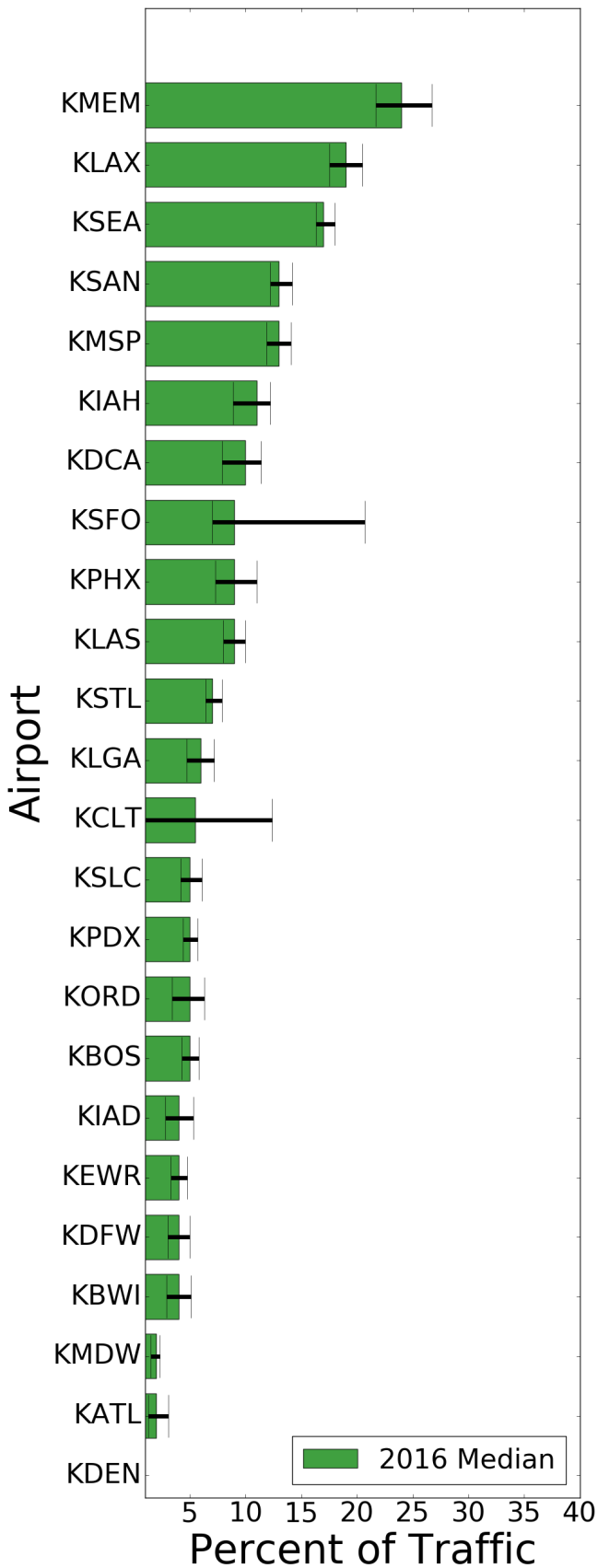
Figure 2. Shows the median adherence percentages for *full-lateral* and *full-vertical* by airport for 2016.
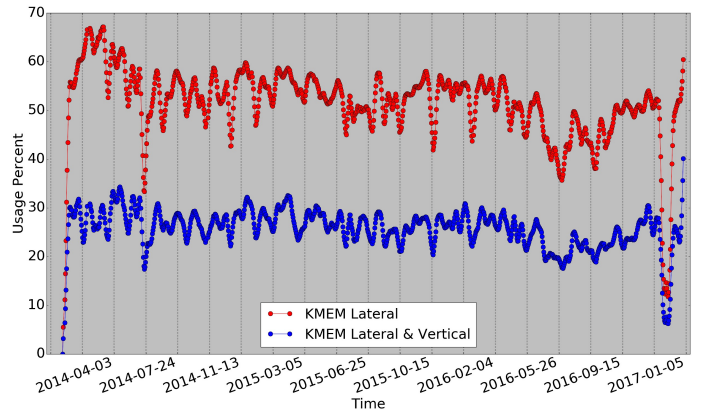

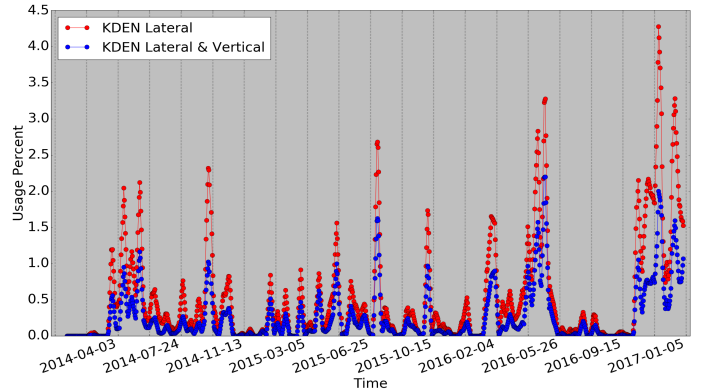Figure 3. Time series of *full-lateral* and *full-later* and *full-vertical* for KMEM.

Figure 4. Time series of *full-lateral* and *full-later* and *full-vertical* for KDEN.

B. Skipped Transition Waypoints

Waypoint skips ranged from nearly 0% to nearly 100%. Greater than 80% of the waypoints in the data-set had skip rates under 20%. Figure 5 shows a histogram of the skip rates across the waypoints with more than 1,000 flight samples over the three years across the 24 airports. To visualize the skipping behavior, we ranked the top 10 most-skipped waypoints along with their flight counts and airports served (see Figure 6). We chose ODF VOR with 53,796 flights and a skip-rate of 89% to visualize on a map. The most frequently used transition path that skipped ODF was the MOL.DIRTY3.RW27L route (FLCON shares the same transition path). This accounted for 14,631 out of the 53,796 ODF skipped flights. These flights were plotted from a top down view alongside the route's transition and can be seen in Figure 7. By visually inspecting the map it would appear the flights are skipping several waypoints after AVERY and getting directly routed to DIRTY.
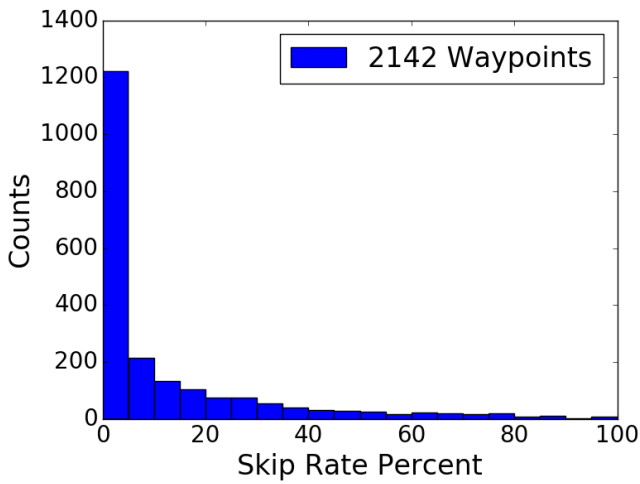
Figure 5. Histogram of the skip percentage rates for all waypoints.

Waypoint Name	Percent Skipped	Airport	Number of Flights
ISSAR	97.14	KLAS	1890
IKRIY	94.18	KLAS	2318
CHEAP	92.78	KIAH	5952
GOTNE	91.12	KMDW	1994
BAMAA	89.13	KLAS	10045
ODF	89.08	KATL	53796
SARAA	85.17	KIAD	1942
JETTZ	84.76	KIAD	1942
FLCON	83.50	KATL	53796
CATTT	82.36	KIAH	11090

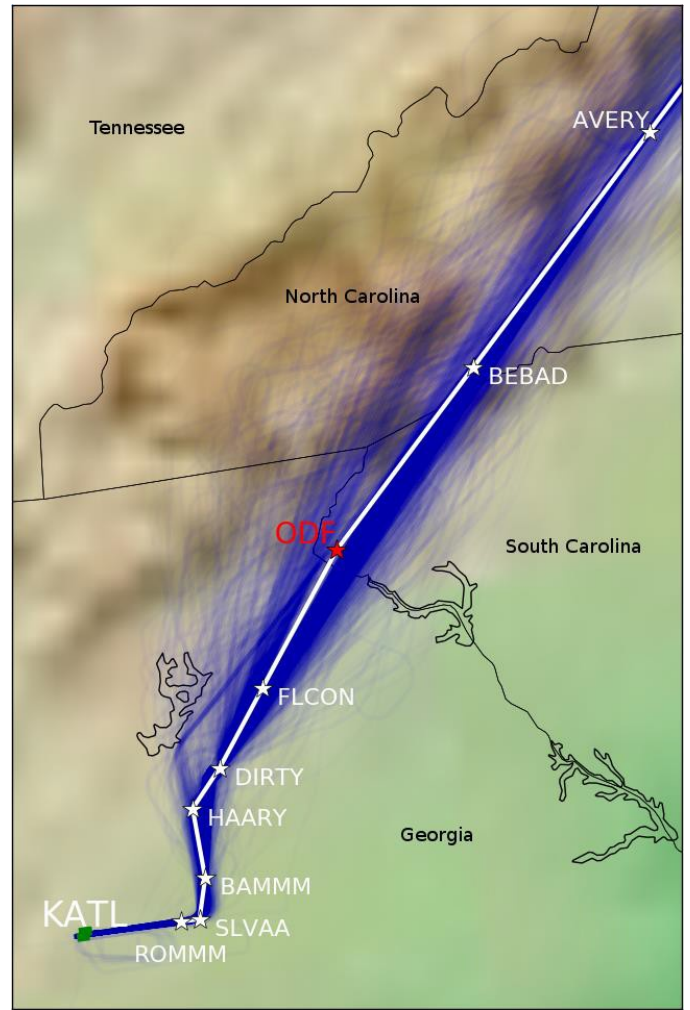Figure 6. The top 10 high skip percentage waypoints with flight counts and airport served.

Figure 7. A top-down map of the 14,631 flights that skipped ODF landing at KATL on the MOL.DIRTY3.RW27L transition.

C. Vertical Characteristics

We looked at the altitudes of aircraft crossing waypoints and two different types of distributions were found for *at* and *window* restrictions. First, we noted that some waypoints had somewhat "normal" distributions centered on the waypoint's altitude restriction. Figure 8 shows a histogram of the altitudes binned at 100 ft. around the *at* type waypoint VASHN. The region between -300 and 300 represents the range where flights are considered compliant with the altitude restrictions and have been removed to focus on the excursion behavior. Even though some of these waypoints showed some skewing, the bulk of the altitudes were clustered around the waypoint restriction. Second, we noticed there were waypoints that had multi-modal distributions located at 1,000 ft. intervals from the waypoint's restriction. Figure 9 shows the histogram binned at 200 ft. below the altitude restriction at the *at* type waypoint JOBEE and illustrates this multi-modal behavior.
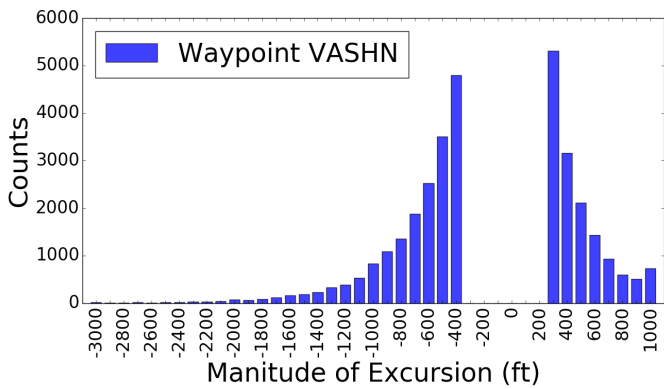
6

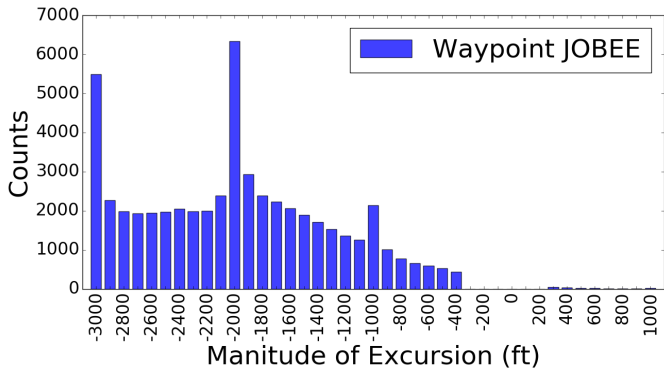Figure 8. Histogram of the excursions for VASHN altitude restriction.

Figure 9. Histogram of excursions for JOBEE's altitude restriction.

Waypoint Type: We looked at aggregate data for the four [waypo]ints: *at*, *ceiling*, *floor*, and *window*. Nonadherence was greatest for [bo]th *at* altitude restrictions. Aircraft flew below these restriction[s] [16%] of the time and 11% above. *Floor* waypoints (at or above) were the [most] vertically missed at three percent below. *Ceiling* (at or below) [we]re missed above at two percent of the time. *Window* (between) [w]ere the most complied with waypoint type with misses either side [arou]nd two percent (see Figure 10).

IV. DISCUSSION

A. Summary

We discovered that the adherence rates for *full-lateral/full-vertical* ranged from 0% percent to 24%. *Full-lateral/full-vertical* was generally less frequent that *full-lateral* only. Differences in rates were not always consistent between *Full-lateral/full-vertical* and f*ull-lateral* only. Waypoint skip rates ranged from nearly 0% to nearly 100%. Greater than 80% of the waypoints in the NAS had skip rates below 20%. Vertical misses occurred 16% and 11% below and above respectively for *at* waypoints. *Window* waypoints were missed between 1% and 2% of the time. *Ceiling* waypoints were missed above the restriction by 2% and *floor* waypoints missed below the restriction 3% of the time. Distributions of altitude misses around waypoints were found to be either normally or multi-

modal. Multi-modal distributions had modes that often occurred in 1,000 ft. increments from the waypoint restriction.

B. Overall Adherence

Characterizing adherence rates is useful for understanding flexibility in the system versus the procedure. Given the immense complexity and quantity of variables associated with the usability of a procedure, low adherence numbers may indicate variables prohibiting use and rendering the procedures unusable. Although, even if utilization is low, aircraft still arrive at the airport, indicating human (ATC) intervention. Some possible explanations for low adherence could be weather, traffic flow rates, ATC policies, procedure design, pilot preferences, and controller preferences.

The NAS currently requires human ATC intervention in many areas to function. However, NextGen is moving toward a pre-planned procedure-based model that could further reduce flexibility in the system. If some of the previously mentioned variables reduce adherence levels, it tells us the procedures have implicit brittleness. Specific variables that contribute to procedural brittleness need to be identified before adding more complexity to the procedures.

C. Adherance Rates Over Time

It is logical that more precise flight paths (e.g., *full-vertical/full-lateral*) would occur less frequently than less precise paths (*full-lateral* only). As procedures are introduced or amended over time, their characteristics may play a role in the adherence rates, which can be monitored by this tool. Further, the differential between the two measurements could indicate changes that were undetectable before and may reveal unexpected future impacts on the procedures.
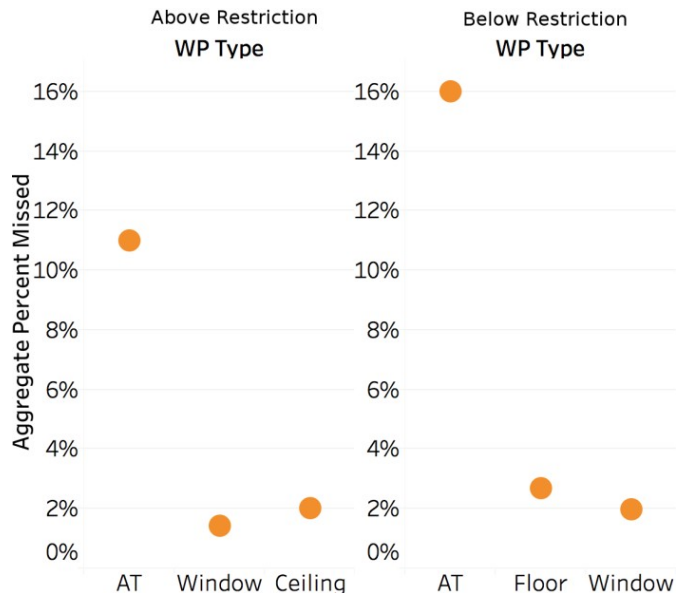

Figure 10. Percent missed below and above the altitude restriction for each waypoint type.

D. Waypoint Skipping

Waypoints skipping behavior could be a useful metric in judging procedure design. In our ASRS review, being removed from a STAR was a primary contributor to deviations from assigned parameters. Since the qualifying criteria for a skip is to enter a transition and depart a transition, a skip in the middle of the procedure would either be a navigation error introduced by reprograming the FMS or ATC intervention, which could be indicative of a design deficiency. Although they have different implications, both are meaningful and could expose problematic portions of STARs.

We noticed a wide disparity in the skip-rates of certain waypoints. Rarely skipped waypoints are likely integral to the specific airspace in which they reside. Conversely, waypoints that are skipped often might be unnecessary or detrimental to normal operations. To fully understand these phenomena, operational personnel would likely need to be interviewed.

E. Vertical Adherence

By comparing different waypoint types, we can possibly see how vertical precision relates to adherence. Although we were not performing inferential statistics in this study, *at* waypoints appeared to be the least complied with restrictions type. This could be related to a greater precision requirement to comply. Further investigation should be done without using aggregate data due to the variation in performance between different STARs and airports.

By looking further into how aircraft cross specific waypoints, the difference between some *multi-modal* and *normal* distributions gives us insight into the operational behavior. Although not completely "normal" these distributions may show a natural excursion behavior of aircraft at specific waypoints. Alternatively, *multi-modal* distributions might show recurring ATC interventions that are occurring systematically. These modes that are situated in 1,000 ft. increments could be the result of ATC assignments. It would be useful to differentiate the two types and investigate them separately

F. Implications

This work is a preliminary investigation and a demonstration of a unique way to view procedural adherence. By utilizing these data, we can see how different airports and ATC facilities utilize procedures. In addition, we can see how certain intrinsic characteristics of procedures compare against each other (e.g., waypoint types). Our ability to mine large amounts of data gives us a look at the airspace system at a meta-level.

As our analysis methods progress we can begin to interface with operational personnel (e.g., controllers and pilots) and add possible explanations to trends found in the data. We believe that this cooperative interdisciplinary approach will eventually yield the most impact for reducing risk and increasing efficiency in designing future procedures.

G. Limitations

1) Intent: We realize trying to quantify pilot-intent without knowing what ATC instructions were issued could be error prone. Flight plans are available and may provide some insight into intent from the planning prospective; however, the filed plans do not always match with the flown procedure. Given this discrepancy we decided that mapping the flight track to the procedure would provide a better measure of the procedure that was most likely implement. In addition, aircraft that are

level and not crossing waypoint altitudes as published are likely not intending to follow those altitudes. Autopilot modes and observed performance could also be a measure of intent that would be applicable to our investigation. Exploring ways to connect additional information sources to enable greater understanding for future experiments would be a necessary step.

2) Causation: Since our data describe behavior of air traffic and are not experimental we cannot make causal inferences. However, we are working to identify relationships between variables to attempt prediction of performance.

3) Measuring Extraneous Variables: This domain is complicated and involves several different user groups (e.g., pilots and controllers), weather, airplanes, automation, and many others variables. Even if relationships between variables are strong we will have challenges generalizing the results to other geographical areas. In addition, we cannot measure the effects of human intervention, which are likely to play a significant role in operations.

H. Future Work

This work has the potential to support several aspects of assessing the RNAV STARs throughout the NAS. One facet is to understand the underlying factors that influence adherence rates. Better understanding can help with future procedure design. A tool to automatically analyze and provide visualization capabilities for otherwise unexplored data would be a valuable decision support tool and aid the procedure design process.

Another use-case is identifying potential risks in operations. This could include analyzing altitude excursions around departure and arrival crossings where potential loss of separation or other risks could occur. Working with airlines to identify other areas of risks would add benefits.

In expanding the tool, additional metrics can be derived from the existing or new features to objectively identify risks and efficiency of these arrivals. This may include:

1. Merging other data sources such as adverse weather forecasts/observations to understand the impact they have on the arrivals or when a procedure has become unusable.
2. Investigating the correlational relationships between variables as they relate to performance, as well as the predictability of aircraft paths using objective data.
3. Leveraging higher sampling rate radar data or and ADSB data that may have more reliable coverage and precision.
4. Integrating flight operational quality assurance data from the airlines where airspeed, autopilot modes, control surfaces, and fuel usage can be used to assess the efficiency of the routes and view pilot behavior.
5. Expanding this investigation to other countries to determine if differences in procedural utilization are seen.

I. Conclusion

We discovered fundamental performance behavior about RNAV STARs in this first exploratory study. This suggests our method of using procedures as the units of measurement is productive. Continuing maturation of this work can help us to gain a better understanding of how procedures are performing so we can improve future designs, help airlines identify risks, and monitor trends for changes in adherence rates. We aim to combine a variety of information sources together in order to achieve these goals. Further, we plan to quantify some of the

factors that add complexity to the air traffic system. We hope our work will increase our understanding of the NAS and knowledge of this complex system.

REFERENCES

[1] Federal Aviation Administration, "Perfromance Based Navigation PBN-NAS Navigation Strategy," 2016

[2] International Federation of Air Line Pilots' Associations "FAA Suspends OPD Arrivals for Atlanta International Airport," 5 Jan 2017.

[3] Dodd, Allen "Boston RNAV STARs" jetBlue Safety, Apr. 2015

[4] Barhydt, Richard, and Catherine A. Adams. *Human factors considerations for performance-based navigation*. National Aeronautics and Space Administration, Langley Research Center, 2006.

[5] Smith, Elida C. "Impact of RNAV terminal procedures on controller workload." Digital Avionics Systems Conference, 2005. DASC 2005. The 24th. Vol. 1. IEEE, 2005.

[6] Metzger, Ulla, and Raja Parasuraman. "The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 43.4 (2001): 519-528.

[7] W. A. Olson and J. E. Olszta, "TCAS Operational Performance Assessment in the U.S. National Airspace," 29th Digital Avionics Systems Conference, Salt Lake City, UT, 2010, pp. 4.A.2-1-4.A.2-11.

doi: 10.1109/DASC.2010.5655351

[8] Erzberger, H.; Davis, T.J.; and Green, S.M.: Design of Center-TRACON Automation System. Proceedings of the AGARD Guidance and Control Panel 56th Symposium on Machine Intelligence in Air Traffic management, Berlin, Germany, 1993, pp. 52-1-52-14.

AUTHORS

Bryan Matthews received his Bachelor's Degree in Electrical Engineering from Santa Clara University in 2002. He has worked at NASA Ames Research Center since 2001 where he is currently a member of the Data Sciences group. His research involves utilizing advanced algorithms to intelligently mine heterogeneous data sources and address complex problems in the national airspace.

Michael Stewart received his MS in human factors from San José State University. He is a senior research associate for the San José State University Research Foundation in the Aerospace Cognitive Engineering group, NASA Ames Research Center. His research has focused on human performance issues with instrument procedures and assessment methods in safety critical operations. His flying background consists primarily of Part 121 airline operations.