CONFLICT RESOLUTION PERFORMANCE IN AN EXPERIMENTAL STUDY OF EN ROUTE FREE MANEUVERING OPERATIONS

Nathan A. Doble, Titan Corporation, nathan.doble@titan.com, Hampton, VA
Richard Barhydt, NASA Langley Research Center, richard.barhydt@nasa.gov, Hampton, VA
James M. Hitt II, Booz Allen Hamilton, hitt_james@bah.com, McLean, VA

Abstract
NASA has developed a far-term air traffic management concept, termed Distributed Air/Ground Traffic Management (DAG-TM). One component of DAG-TM, En Route Free Maneuvering, allows properly trained flight crews of equipped “autonomous” aircraft to assume responsibility for separation from other autonomous aircraft and from Instrument Flight Rules (IFR) aircraft. Ground-based air traffic controllers continue to separate IFR traffic and issue flow management constraints to all aircraft.

To examine En Route Free Maneuvering operations, a joint human-in-the-loop experiment was conducted in summer 2004 at the NASA Ames and Langley Research Centers. Test subject pilots used desktop flight simulators to resolve traffic conflicts and adhere to air traffic flow constraints issued by subject controllers. The experimental airspace integrated both autonomous and IFR aircraft at varying traffic densities.

This paper presents a subset of the En Route Free Maneuvering experimental results, focusing on airborne and ground-based conflict resolution, and the effects of increased traffic levels on the ability of pilots and air traffic controllers to perform this task. The results show that, in general, increases in autonomous traffic do not significantly impact conflict resolution performance. In addition, pilot acceptability of autonomous operations remains high throughout the range of traffic densities studied. Together with previously reported findings, these results continue to support the feasibility of the En Route Free Maneuvering component of DAG-TM.

Introduction
Although air traffic growth has slowed in recent years, demand for air travel is once again rising to levels that will challenge airspace system capacities in the United States and Europe [1,2]. Impending air traffic controller retirements will place further strains on the system [3]. As a result, it has been realized that a transformational, rather than evolutionary, approach to air traffic control modernization is needed to handle future airspace demands [4].

As part of the Advanced Air Transportation Technologies (AATT) project, NASA has developed such a far-term, transformational concept for air traffic management, called Distributed Air/Ground Traffic Management (DAG-TM) [5]. DAG-TM seeks to increase capacity and maintain safety by redistributing decision-making responsibility among airborne and ground-based elements of the air transportation system. It is a gate-to-gate concept, addressing all flight phases from pushback to arrival.

En Route Free Maneuvering
En Route Free Maneuvering is one component of DAG-TM, addressing operations in the en route and terminal-transition phases of flight [6]. Under the En Route Free Maneuvering paradigm, properly trained crews of properly equipped aircraft assume responsibility for traffic separation. In return for complying with basic traffic flow management constraints (e.g., meeting a required crossing time at a terminal area arrival waypoint), such flight crews are free to modify their path in real time, without approval from a ground-based Air Traffic Service Provider (ATSP). These flights would operate under a new set of flight rules called Autonomous Flight Rules (AFR).

Except for busy terminal areas, where AFR operations would not be allowed, AFR traffic would operate in the same airspace as Instrument Flight Rules (IFR) and Visual Flight Rules (VFR) aircraft. For those operators who choose not to equip for AFR, the ATSP would continue to provide separation services. By distributing separation assurance among multiple airborne and ground-based elements, the National Airspace System may be able to accommodate a higher increase in demand beyond what is possible with a centralized, ground-based system.
The Autonomous Operations Planner

Central to AFR operations are the capabilities of airborne conflict prevention, detection, and resolution. It is assumed that a pilot will not be able to safely perform these functions without some form of conflict alerting and decision support. As such, NASA Langley Research Center has developed a prototype airborne toolset, called the Autonomous Operations Planner (AOP), to assist the pilot in these tasks [7,8].

The prototype AOP interface is designed around a modern “glass cockpit” flight deck. It provides conflict alerting and resolution advisories via the navigation display, using state and intent information from the ownship and traffic aircraft. Additional AOP functions are accessed via the Flight Management System (FMS) control display unit. Depending on the mode of aircraft control and the proximity of the conflict, AOP presents either strategic resolution advisories (modifications to the FMS active route) or tactical resolution advisories (simple heading or vertical speed command “bugs”). In this way, AOP supports the pilot in all flight modes [9]. In addition, to avoid the inadvertent creation of new traffic conflicts while maneuvering, AOP displays conflict prevention information in the form of Maneuver Restriction Bands—“no-fly” regions on the heading and vertical speed displays. Figure 1 shows an example of the AOP interface with a strategic resolution advisory and Maneuver Restriction Band displayed.

Figure 1. AOP Information on Navigation Display, Including Strategic Resolution Advisory

Previous Work

The work presented in this paper builds upon past DAG-TM research conducted at NASA as well as initial Free Flight research by organizations such as NLR in the Netherlands [10]. As reported at earlier USA/Europe Air Traffic Management R&D Seminars, previous NASA experiments included research on over-constrained conflict scenarios, AFR operations in confined airspace, the ability of AFR pilots to regain lost separation, and the effect of intent on decision making [11-14].

The remainder of this paper discusses the design and a subset of the results of an experiment evaluating AFR operations with the AOP toolset. This experiment extended the previous research in several significant ways. First, it included a realistic, mixed-equipage operating environment, containing both AFR and IFR flights. Arrival profiles were simulated as well as overflights, and the functionality of AOP was enhanced to provide vertical resolution guidance in addition to lateral guidance. Interactions with a ground-based ATSP were also modeled with the addition of air traffic controller test subjects. Finally, multiple traffic densities were tested such that the potential scalability of AFR operations could be evaluated.

Experimental Approach

Research Issues

In summer 2004, NASA Langley and Ames Research Centers conducted a joint, human-in-the-loop simulation of AFR operations. The experiment investigated the following two research issues:

- **Mixed Operations**: How does the safety and efficiency of mixed operations (i.e., both AFR and IFR aircraft in the same airspace) compare to ATSP-managed operations (i.e., all-IFR) at traffic densities near current-day en route sector capacities?
- **Scalability**: What is the feasibility of further increasing the total number of aircraft in a sector (beyond what the ATSP could normally control), if the number of IFR aircraft remains at or below current-day levels?

This paper primarily addresses the issue of scalability, specifically in the context of resolution maneuvers for scripted traffic conflicts. Mixed operations will be addressed further in a future publication.
Participants

Test subjects included 12 pilots at NASA Langley as well as pilots and air traffic controllers at NASA Ames. The NASA Langley subject pilots were all active, furloughed, or recently retired pilots with Airline Transport Pilot ratings and experience in Boeing glass cockpit aircraft.

During the experiment, the NASA Langley pilots flew workstation-based flight simulators that emulated the displays of a Boeing 777, with the inclusion of an AOP. Additional AFR and IFR background traffic was supplied using multi-aircraft pseudo-pilot stations at both NASA Ames and NASA Langley. These stations were manned by research personnel, acting as confederate pilots.

As shown in Figure 2, the experimental airspace was comprised of simulated high- and low-altitude sectors in the Fort Worth Air Route Traffic Control Center, and scenarios were designed to approximate realistic traffic flows into, out of, and over the Dallas-Ft. Worth area. The sectors were staffed at NASA Ames by 5 professional air traffic controller test subjects. The controllers provided separation services only for IFR-IFR traffic conflicts. Additionally, researchers acted as pseudo-controllers in large “ghost” sectors that surrounded the experimental sectors, providing limited services to aircraft entering and exiting the subject-controlled airspace.

Figure 2. DAG-TM Experimental Airspace

Laboratories at Ames and Langley were linked via a dedicated internet connection, allowing Automatic Dependent Surveillance-Broadcast (ADS-B), Controller-Pilot Data Link Communications (CPDLC), and radio communications to all be modeled in the experiment.

Scenario Design

The experiment was designed in a within-subjects format, with subject pilots flying 16 different scenarios. As the independent variable, four different traffic conditions were simulated, shown in Table 1. Note that the traffic density increases were solely due to increases in overflight traffic. Arrival flows were kept at a constant, near-saturated level throughout the experiment.

### Table 1. Traffic Conditions Tested

<table>
<thead>
<tr>
<th>Condition</th>
<th>Avg. Traffic Density</th>
<th>% IFR</th>
<th>% AFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>slightly above current Monitor Alert parameter</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>C2</td>
<td>equal to C1 density</td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>C3</td>
<td>≈1.5 × C1 density</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>C4</td>
<td>≈2 × C1 density</td>
<td>35%</td>
<td>65%</td>
</tr>
</tbody>
</table>

At each of the traffic conditions, pilots flew arrival profiles (starting at cruise and ending at a terminal airspace arrival fix) in two scenarios and overflight profiles in two scenarios. Subject pilots flew under AFR for all scenarios except those in C1. When operating under AFR, subject pilots were responsible for resolving any traffic conflicts that arose, both with other AFR aircraft and with IFR aircraft. For arrival profiles, subject pilots were also responsible for meeting a required crossing time, altitude, and airspeed at a terminal area arrival fix.

Scripted Traffic Conflicts

In order to have a set of comparable traffic conflicts for analysis, carefully constructed conflicts were deliberately scripted to occur in each scenario. For all C1 scenarios, six “controller-focused” IFR-IFR conflicts were scripted for the subject controller to resolve (one per NASA Langley subject-piloted IFR overflight). For scenarios at C2, C3, and C4, six “pilot-focused” AFR-IFR conflicts were scripted for the subject pilots to resolve (one per Langley AFR overflight). In addition, every scenario at C2, C3, and C4 contained one additional controller-focused IFR-IFR scripted conflict (not involving a subject pilot) for the controller to resolve. This was done so that groundside-resolved conflicts could be compared across traffic levels in a similar manner as airborne-resolved conflicts.

The conflicts were all constructed using a similar geometry, with a co-altitude intruder at a 60-degree convergence angle, approaching either from the left or right. In addition to scaling overflight traffic as a whole, traffic was locally scaled in a structured manner around each scripted conflict.
through the addition of proximate but non-conflicting “flanking” aircraft. For C1 and C2 scenarios, one flanking aircraft was added at -1000 ft relative altitude. At C3, a second, opposite-direction aircraft was added at +1000 ft relative altitude. At C4, a third, co-altitude aircraft was added. Figure 3 shows this scripted conflict geometry.

Figure 3. Scripted Conflict Geometry

For pilot-focused conflicts, the intruder aircraft was IFR and all other aircraft were AFR. For controller-focused conflicts in traffic conditions C2, C3, and C4, the conflicting aircraft were both IFR and the flanking aircraft were all AFR. By definition, all controller-focused scripted conflict aircraft in C1 scenarios were IFR.

Subject pilots were not informed that each scenario contained scripted conflicts. To avoid unwanted practice effects and to prevent subject pilots from identifying scripted conflicts, several precautions were taken to minimize the repetitiousness of the conflicts. Headings, cruise altitudes, call signs, and conflict times (relative to scenario start) were all varied across scenarios. Post-experiment debriefs indicated that these strategies were successful in masking the scripted nature of these conflicts.

Results

Sixty-nine out of a possible 72 pilot-focused scripted conflicts (6 AFR overflights in each of 12 scenarios at C2, C3, and C4) occurred. Results for these conflicts are discussed below in the Airborne Objective Metrics and Airborne Subjective Metrics sections. While each subject pilot’s trajectory was designed to be conflict-free until the scripted conflict, unpredictable traffic interactions (e.g., other traffic conflicting with the subject pilot due to resolution maneuvers earlier in the scenario) prevented some scripted conflicts from taking place.

Thirty-one out of a possible 36 controller-focused scripted conflicts (6 in each of 4 scenarios at C1, plus 1 in each of the 12 remaining scenarios) occurred. Results for these conflicts are discussed in the Groundside Metrics section.

Airborne Objective Metrics

Objective data were used to assess the effects of traffic density on the following AFR performance measures: time in conflict, time before Loss of Separation (LOS) at which conflicts are resolved, the frequency of induced conflicts, and the frequency of multiple resolutions.

Time in conflict was used as a measure of the difficulty of resolving a conflict. Notionally, more difficult conflicts should result in longer AOP resolution advisory computation times and longer pilot decision-making times. Figure 4 shows mean conflict durations for pilot-focused scripted conflicts at each traffic condition (C2-C4). Error bars represent one standard deviation. To test if significant differences in conflict duration times existed across traffic conditions, a within-subjects ANOVA was performed. The results showed slight increases in mean conflict duration as traffic was increased, but these differences were not significant ($F(2,22) = 0.89, p > 0.05$).

Figure 4. Conflict Duration vs. Traffic Condition

Time before LOS at conflict resolution was used as a measure of safety and resolution urgency. For the purposes of this experiment, LOS was defined as two aircraft within 5 nm laterally and 950 ft vertically\(^1\). The AOP nominally provided a ten-minute lookahead time for conflict alerting, unless restricted by ADS-B

---

\(^1\) The vertical threshold was set at 950 ft instead of 1000 ft to reduce the number of false alerts resulting from small flight path perturbations for aircraft at adjacent flight levels.
range limitations or data losses. Figure 5 shows mean times to LOS at conflict resolution at each traffic condition, with error bars representing one standard deviation. To test if significant differences existed across traffic conditions, a within-subjects ANOVA was performed. While mean times to LOS slightly decreased as traffic levels increased, these differences were not significant \((F(2,22) = 0.82, p > 0.05)\).

![Figure 5. Time to LOS at Conflict Resolution vs. Traffic Condition](image)

The frequency of induced conflicts was used as a measure of system stability. Induced conflicts were defined as additional (non-scripted) conflicts caused by a previous resolution maneuver. Of 69 scripted conflicts across traffic conditions C2, C3, and C4, only 7 caused induced conflicts. All of these occurred in C4.

The frequency of multiple resolutions was also a measure of system stability. If a subject pilot was in conflict with the same intruder multiple times, and the subject pilot was required to make an additional resolution maneuver, this was noted as a multiple resolution conflict. Figure 6 shows the percentage of scripted conflicts that required multiple resolutions at each traffic condition. To test if significant differences in the frequency of multiple resolutions existed across traffic conditions, a \(\chi^2\) test was performed. The results showed a slight increase in the mean percentage of multiple resolutions as traffic levels increased. However, these differences were not significant \((\chi^2(2, N = 69) = 0.41, p > 0.05)\).

![Figure 6. Multiple Resolutions vs. Traffic Level](image)

**Airborne Subjective Metrics**

Pilot-reported subjective data were used to assess the effects of traffic density on workload, safety, and acceptability of AFR operations. These results were recorded using written questionnaires at the conclusion of each scenario. While these results address scenario-wide impressions, they are considered pertinent to the discussion, as resolution of scripted conflicts was a major component of overflight scenarios.

To examine workload, pilots were asked “How would you rate your overall workload in this scenario?” on a scale from 1=completely acceptable to 7=completely unacceptable. The 12 subject pilots provided a total of 84 responses. Figure 7 shows mean workload ratings at each traffic condition, with error bars representing one standard deviation. A within-subjects ANOVA was performed to test for significant differences in workload ratings across traffic conditions. The results indicated that no significant differences existed across the three traffic conditions \((F(2,22) = 3.24, p > 0.05)\). However, it is worth noting that this traffic condition comparison was nearly significant \((p = 0.059)\).

![Figure 7. Pilot-Reported Workload vs. Traffic Condition](image)

To examine safety, subject pilots were asked “Considering the complete start-to-end scenario (including conflicts and your resolution actions), what was the level of safety?” Pilots rated safety on a scale from 1=completely unsafe to 7=completely safe. The 12 subject pilots provided a total of 84 responses. Figure 8 shows mean safety ratings and standard deviations at each traffic condition. A within-subjects ANOVA showed no significant differences in safety ratings across traffic conditions \((F(2,22) = 0.62, p > 0.05)\).
To examine the acceptability of AFR operations, subject pilots were asked “Based on your overall workload during this scenario and considering other tasks you normally perform during line operations that are not being simulated, how would you estimate the acceptability of AFR operations during a typical flight?” Pilots rated acceptability on a scale of 1=completely unacceptable to 7=completely acceptable. The 12 subject pilots provided a total of 84 responses. Figure 9 shows mean acceptability ratings and standard deviations at each traffic condition. While mean acceptability ratings slightly decreased as traffic levels increased, a within-subjects ANOVA indicated no significant differences in acceptability across traffic conditions ($F(2,22) = 1.87, p > 0.05$).

**Groundside Metrics**

Similar to the airborne metrics, groundside data were recorded for the controller-focused IFR-IFR scripted conflicts for metrics of duration, time at resolution, induced conflicts (none occurred), multiple resolutions (none occurred), and subjective workload. However, due to the small number of data points in the mixed-traffic scenarios (3 at C2, 3 at C3, and 2 at C4 vs. 23 at C1), statistical analyses were not performed on these data. Rather, histograms are provided to qualitatively assess each metric. And while C1 scripted conflicts cannot be evaluated directly against conflicts at C2, C3, and C4 (due to different flight rules for the flanking aircraft), means and standard deviations at C1 are shown on each chart for comparison.

Figure 10 shows durations for the controller-focused scripted conflicts. No consistent trend is observed across the traffic levels and durations in mixed-traffic scenarios are comparable to durations at C1.

Figure 11 shows times to LOS at conflict resolution for the controller-focused scripted conflicts. No consistent trend is observed across traffic conditions and resolution times are comparable to times at C1. Unlike the airborne-resolved conflicts, resolution times for controller-resolved conflicts are somewhat dependent on airspace geometry. Whereas AOP provided a 10 minute alerting lookahead window regardless of sector boundaries, controller tools were limited to alerting for aircraft within a particular airspace sector.

Figure 12 shows controller-reported workload values, on a 1 to 5 scale (5=highest workload). Unlike the Langley pilots, controllers reported instantaneous workload throughout the
duration of each scenario. The values used in Figure 12 are those nearest to the time when the conflict was resolved. No consistent trend is observed across traffic conditions, but values for mixed-traffic scenarios are generally lower than values at C1.

![](image)

**Figure 12. Controller-Reported Workload**

**Discussion**

The above results can be used to answer two high-level questions concerning the feasibility of integrated AFR-IFR operations: 1) How does increased AFR traffic affect a pilot's ability to solve AFR-IFR conflicts? and 2) How does increased AFR traffic affect a controller’s ability to solve IFR-IFR conflicts?

**Airborne Conflict Resolution**

In general, increases in traffic density caused no significant impacts to airborne conflict resolution performance. As expected, conflict durations increased slightly with traffic increases, but these differences were not significant. Mean conflict durations of approximately two minutes were observed, which compares favorably with durations of controller-resolved conflicts. Likewise, resolution times relative to LOS decreased slightly as traffic increased, but these differences were also not significant. Moreover, mean resolution times were well greater than 5 minutes, the time at which the AOP alert level increases and pilots are required under AFR procedures to take swift resolution action.

It is a concern that seven induced conflicts were observed at C4 while none were observed at C2 or C3. Further experimentation is needed to accurately determine the effect of traffic density on induced conflicts. However, it should be noted that an induced conflict rate of 29% (7 conflicts induced by 24 scripted conflict resolutions at C4) is not unusual, as compared to previous results [12,14]. Moreover, all but one of these induced conflicts was caused by a manual maneuver, for which the pilot did not seek or use AOP resolution guidance (e.g., manual FMS path modifications, vertical speed changes without reference to AOP tactical advisory bugs). Thus, the incidence of induced conflicts may be mitigated through more extensive pilot training and modifications to AFR procedures. Human factors issues have also been identified in the AOP conflict prevention tools, which are designed to alert pilots to potential conflicts along proposed manual maneuvers. For example, in the version of AOP used during the experiment, it was possible for conflict prevention symbology to “time out” on the aircraft displays, thus masking the existence of potential conflicts.

The frequency of conflicts requiring multiple resolutions slightly increased with increases in traffic, but these differences were not significant. Upon review of the resolution strategies for these conflicts, no obvious correlation was seen between the type of resolution (strategic or tactical, lateral or vertical) and the occurrence of multiple resolutions. Improvements to AOP logic, such as the addition of spatial or temporal buffers for resolution maneuvers, may reduce the incidence of conflicts requiring multiple resolutions. No such buffers were used for this experiment.

Pilot-reported subjective responses were generally quite supportive of AFR operations, even in high-density scenarios. No consistent trends were observed across traffic conditions for workload or safety. At C4 (not unexpectedly having the highest workload value), mean workload was still relatively low, at 2.1 on a 1 to 7 scale. While these results are encouraging, full-scale aircraft simulations including two-pilot operations would be necessary to thoroughly evaluate workload. Finally, while pilot-reported AFR acceptability ratings decreased slightly with increases in traffic, these differences were not significant. Even at C4, only 2 of 29 responses were below 5 on a 1 to 7 acceptability scale.

**Groundside Conflict Resolution**

Due to limited data, less can be concluded from the controller-resolved conflicts. However, the addition of AFR traffic (beyond levels that a controller could handle were all aircraft IFR) seems to have little impact on the ability of controllers to solve IFR-IFR traffic conflicts.

For mixed-traffic conditions C2, C3, and C4, results for conflict duration and time to LOS at resolution were comparable to results at the all-IFR condition C1. Moreover, no obvious trends were observed across traffic conditions. The absence of any induced conflicts or multiple resolutions also
supports feasibility of mixed AFR-IFR operations at high traffic densities. Finally, controller-reported instantaneous workload values were actually lower, in general, for the mixed-traffic conditions than for the all-IFR condition. This is not unsurprising, as conditions C2, C3, and C4 all contained a level of IFR traffic lower than at C1. Thus, these results support the theory that AFR traffic contributes less to controller workload than an equivalent amount of IFR traffic. However, further research and more extensive data is required to draw stronger conclusions about the performance of controllers in a high-density, mixed AFR-IFR environment.

Conclusions

The DAG-TM experiment conducted at NASA Ames and Langley Research Centers has extended previous research findings by examining mixed-equipage, En Route Free Maneuvering operations in the high-density traffic environment that is predicted to emerge in the coming decades. Neither airborne nor ground-based conflict resolution performance was found to significantly degrade as traffic was increased. Pilot-reported subjective ratings for workload, safety, and AFR acceptability remained favorable throughout the range of traffic conditions studied. In addition, as compared to results in all-IFR scenarios, controller workload remained relatively low while resolving conflicts in mixed-traffic conditions. These findings further support the feasibility of the En Route Free Maneuvering component of DAG-TM.

Acknowledgements

The authors would like to thank Paul Lee at NASA Ames for compiling data for the groundside metrics.

References


Key Words
Autonomous, Conflict Resolution, DAG-TM, En Route, Free Flight, Simulation

Author Biographies
Nathan A. Doble is an ATC Research Scientist in the Air Traffic Systems Division of The Titan Corporation. He holds B.S. and M.S. degrees in Aeronautics and Astronautics from the Massachusetts Institute of Technology, where he conducted research on ATC interfaces for airport control towers. He currently supports air traffic management initiatives at NASA Langley Research Center in the areas of concept development, experiment design and analysis, and human factors. He holds a private pilot certificate and is training for his instrument rating.

Richard Barhydt has a B.S. in Aerospace Engineering from the University of Colorado (1995) and an M.S. in Aeronautics and Astronautics from MIT (1997). He is an aerospace engineer at the NASA Langley Research Center, working on the Advanced Air Transportation Technologies project. His research has included the development of prototype flight deck decision support systems and airborne procedures for NASA’s Distributed Air-Ground Traffic Management concept. He is a certified instrument flight instructor and commercial pilot for single engine aircraft. He is also a member of RTCA Special Committee 186, Working Group 6 (ADS-B MASPS).

James M. Hitt II, an Associate with Booz Allen Hamilton, has more than 8 years of experience in the area of research, development, and assessment of human performance in the aviation domain. Jim has most recently supported numerous government agencies including NASA, FAA, and TSA in aviation-related human factors programs. Mr. Hitt serves as an associate editor for Ergonomics in Design and is an author of over 30 book chapters, journal articles, and conference proceedings papers. He has a M.S. degree in Experimental Psychology and is completing his Ph.D. in Human Factors from the University of Central Florida.