Flight Crew Responses to the Interval Management Alternative Clearances (IMAC) Experiment

Brian T. Baxley\textsuperscript{1}, Sara R. Wilson\textsuperscript{2}, Kurt A. Swieringa\textsuperscript{3}, and Roy D. Roper\textsuperscript{4}

NASA Langley Research Center (LaRC), Hampton VA, 23681

Interval Management Alternative Clearances (IMAC) was a human-in-the-loop simulation experiment conducted to explore the efficacy and acceptability of three IM operations: CAPTURE, CROSS, and MAINTAIN. Two weeks of data collection were conducted, with each week using twelve subject pilots and four subject controllers flying ten high-density arrival scenarios into the Denver International Airport. Overall, both the IM operations and procedures were rated very favorably by the flight crew in terms of acceptability, workload, and pilot head down time. However, several critical issues were identified requiring resolution prior to real-world implementation, including the high frequency of IM speed commands, IM speed commands requiring changes to aircraft configuration, and ambiguous IM cockpit displays that did not trigger the intended pilot reaction. The results from this experiment will be used to prepare for a flight test in 2017, and to support the development of an advanced IM concept of operations by the FAA and aviation industry.

Nomenclature

\begin{center}
\begin{tabular}{ll}
ATD & = Air Traffic Management Technology Demonstration \\
ASTOR & = Aircraft Simulator for Traffic Operations Research \\
ATOL & = Air Traffic Operations Laboratory \\
CGD & = Configurable Glass Display \\
DTS & = Development and Test Simulator \\
EFB & = Electronic Flight Bag \\
IFD & = Integrated Flight Deck \\
IM & = Interval Management \\
IMAC & = Interval Management Alternative Clearances \\
Ownship & = Aircraft with IM equipage \\
PF & = Pilot Flying \\
PM & = Pilot Monitoring \\
Target & = Aircraft the IM aircraft is assigned to follow \\
\end{tabular}
\end{center}

I. Introduction

The goal of the Air Traffic Management Technology Demonstration-1 (ATD-1) Concept of Operations (ConOps) is to increase the throughput at high-density airports while simultaneously increasing the efficiency of the arrival operations\textsuperscript{1}. The ATD-1 ConOps combines advanced arrival scheduling, controller decision support tools, and aircraft avionics to enable multiple time deconflicted, efficient arrival streams into a high-density terminal airspace. Aircraft on these arrival streams primarily use speed control along their optimized profile descents to

\textsuperscript{1} Research Engineer, Crew Systems and Aviation Operations Branch, NASA LaRC, AIAA Senior Member.
\textsuperscript{2} Research Engineer, Systems Engineering and Engineering Methods Branch, NASA LaRC.
\textsuperscript{3} Research Engineer, Crew Systems and Aviation Operations Branch, NASA LaRC, AIAA Member.
\textsuperscript{4} Research Engineer, Crew Systems and Aviation Operations Branch, NASA LaRC.
maintain adequate separation from other aircraft and to achieve precise schedule conformance, decreasing the number of instances that aircraft are vectored off path.

The goal of Interval Management (IM), one of the three ATD-1 technologies, is to increase the precision of spacing between aircraft in order to improve traffic flow and airport throughput, and potentially reduce the overall voice communication requirement for controllers. IM, comprised of the software in the aircraft avionics and the procedures to use that software, is designed to support the ATD-1 ConOps by enabling the enroute controller to issue a single strategic clearance to the flight crew to achieve a specific time or distance behind the preceding aircraft, referred to as the Target aircraft. The IM software onboard the aircraft uses Ownship data (route of flight, current location, etc.), the IM clearance issued by the controller (Target to follow, the spacing interval behind the Target, etc.), and the Target aircraft’s transmitted Automatic Dependent Surveillance–Broadcast (ADS-B) state data to calculate the airspeed necessary for the IM-equipped aircraft to achieve the assigned spacing interval.

The IMAC experiment expanded the ATD-1 ConOps and the IM CROSS procedure from previous research,1,2 by investigating two new IM clearance types: the CAPTURE and the MAINTAIN operation. A brief description of these three IM clearances and operations is given below.

1. CROSS
   The goal of the CROSS clearance is to achieve the Air Traffic Control (ATC) specified assigned spacing goal at a designated achieve-by point, and then maintain that spacing until the planned termination point. The CROSS clearance can be used both when the Ownship and Target aircraft are on different routes and when they are on the same route, and the controller has a specific spacing goal to be met at the achieve-by point. If the achieve-by point and planned termination point coincide, there is no maintain phase of the IM operation.

2. CAPTURE
   The goal of the CAPTURE clearance is to capture an assigned spacing goal then maintain that assigned spacing goal until the planned termination point. The CAPTURE clearance is used when the Ownship and Target aircraft are on the same route, and the controller has a specific assigned spacing goal to be maintained between the two aircraft. There is no achieve-by point since the spacing algorithm will command speeds that result in a moderate rate of closure to the assigned spacing goal interval.

3. MAINTAIN
   The goal of the MAINTAIN clearance is to maintain the current spacing interval between the Ownship and Target aircraft until the planned termination point. Similar to the CAPTURE clearance, the MAINTAIN clearance can only be issued when the Ownship and Target aircraft are on the same route. Unlike the CROSS and CAPTURE clearances that are intended for use during arrival operations when metering is in effect, the MAINTAIN clearance is intended for use during enroute metering operations (for example, miles-in-trail) or when metering is not being used during arrival operations.

The Interval Management Alternative Clearances (IMAC) human-in-the-loop experiment was conducted by the ATD-1 Project in the summer of 2015 at NASA’s Langley Research Center. Six different simulator types at three different locations were connected to each other to emulate as realistically as possible the dynamic and complex high-density arrival operations that occur into major airports. The experiment participants consisted of two groups of twelve subject pilots and four subject controllers. Each pilot group flew ten high-density arrival scenarios into the Denver International Airport. In addition to the subject controllers and pilots, each group also included two confederate controllers who controlled aircraft outside of the experiment test area, and six confederate pilots who helped provide voice communications for the aircraft that were not flown by the subject pilots.

II. Experiment Design

A. Objectives
   The objective of this experiment was to assess the acceptability and system performance of three different types of IM clearances during realistic, high-density arrival operations. In particular, the experiment focused on the acceptability and workload of the IM operations for the flight crew, and the identification of possible issues with implementation into real-world operations.
B. Simulation Environment

The simulation environment was the airspace and instrument arrival procedures into the Denver International Airport. This airport was specifically chosen since all the published Standard Terminal Arrival Routes (STARs) connect either to a Required Navigation Procedure (RNP) or Instrument Landing System (ILS) approach. Each scenario contained aircraft arriving from all four directions, and landing on one of two runways (Figure 1). In the simulation environment, aircraft can be flown 1) by research subjects (the only source for reported data) flying one of three simulator types, 2) by confederate pilots to create additional traffic necessary to create the high density arrival operations the research subjects were flying in, and 3) by a software-based pilot model to visually create complex arrival flows on the opposite side of the airspace and airport that are completely independent from any of the pilot flown aircraft.

For simplicity, subject controlled, confederate controller, and pilot-model controlled simulated aircraft will simply be referred to as “aircraft”, however the data reported is based solely on the subject controlled aircraft.

![Figure 1. Arrival procedures into Denver for north-flow (left panel) and south-flow landing (right panel).](image)

For each scenario, only aircraft in one half of the simulation airspace were flown by subject and confederate pilots and controlled by controllers, while aircraft in the other half of the airspace were flown by an automated pilot model. All aircraft landed on the runway closest to their direction of arrival with human operated aircraft on one runway and automated aircraft on the parallel runway. Figure 2 provides an illustration of a scenario where the subject flown aircraft are in blue arriving from the west, and the green aircraft on the west are flown by the confederate pilots, as are all the orange overflight aircraft. The green aircraft on the east side of the Denver International Airport are all completely automated and did not interact with the controllers nor provide any data reported in this paper (these aircraft were to add visual realism and complexity to controller and cockpit displays only). There were no departures from KDEN in the scenario; however, there were overflight aircraft for additional realism.

In total, there were 229 arrival operations conducted into Denver International Airport that were used for data analysis, of which 113 were IM operations. Shortly after each scenario began, the controller would issue the flight crew of IM equipped aircraft an IM clearance. The clearance type was determined by the scenario, and the assigned spacing goal (ASG), was calculated by the Time Based Flow Management (TBFM) scheduling software. The TBFM calculations used wake vortex separation criteria for operations in poor weather, and therefore, the spacing between aircraft ranged from approximately 70 to 110 seconds. The IM achieve-by point was the final approach fix, as was the planned termination point.
C. Test Matrix

The IMAC experiment utilized the 1x5 test matrix shown in Table 1 to assess the acceptability and performance of three different types of IM clearances. Two replicates of each scenario were conducted – one with the Captain as the pilot monitoring (PM) and the First Officer as the pilot flying (PF), and the other replicate with the roles reversed. Therefore, each flight crew flew a total of ten scenarios during data collection.

Table 1. IMAC Experiment Test Matrix

<table>
<thead>
<tr>
<th>BASELINE</th>
<th>CAPTURE</th>
<th>CROSS</th>
<th>MAINTAIN</th>
<th>MIXED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario A</td>
<td>Scenario C</td>
<td>Scenario E</td>
<td>Scenario G</td>
<td>Scenario I</td>
</tr>
<tr>
<td>Scenario B</td>
<td>Scenario D</td>
<td>Scenario F</td>
<td>Scenario H</td>
<td>Scenario J</td>
</tr>
</tbody>
</table>

A within-subject design was utilized, and the run order for the BASELINE, CAPTURE, CROSS, and MAINTAIN scenarios was partially counterbalanced using a Latin square design. During BASELINE scenarios, the flight crew used current day operations and the IM equipment and procedure were not used. A detailed description of the controller displays and procedures during the experiment is available in Ref. 4.) These eight scenarios were followed by the two MIXED scenarios in which the controllers could issue any of the three IM clearance types.

D. Facilities

The subject controllers operated from the Air Traffic Operations Laboratory (ATOL), which consists of multiple configurable spaces that can be established as enroute, terminal, and tower control facilities, with the appropriate hardware and displays. Both enroute and terminal controller simulators were used during IMAC, with the radar scope displays driven by the Multi-Aircraft Control System software system developed at NASA Ames Research Center.

Half of the subject pilots also operated from the ATOL, and flew dual-crew, three-monitor Aircraft Simulation for Traffic Operations Research (ASTOR) desktop simulators (Figure 3). One quarter of the subject pilots flew the Development and Test Simulator (DTS), a full-mission simulator representative of a large generic commercial transport category aircraft (not shown), while the remaining one quarter flew the Integration Flight Deck (IFD), a full-mission Boeing 737-800 simulator (Figure 4). The location of the electronic flight bag (EFB) and configurable graphics display (CGD) in the ATOL (highlighted by magenta boxes in Figure 3) was positioned as closely as feasible to the IFD (highlighted by magenta boxes in Figure 4). The location of the EFB and CGD in the DTS was practically identical to the IFD.
The two confederate controllers in the ATOL facilitated setting up each scenario and performed tower operations. The six confederate pilots flew single-crew dual-monitor Multi-Aircraft Control System pseudo-pilot stations in the ATOL to create high-density arrival operations with a realistic voice communications environment.

![Figure 3. Aircraft Simulation for Traffic Operations Research (ASTOR) stations.](image)

E. Flight Crew Interfaces for IM Operations

The cockpit human-machine interface devices used during the IMAC experiment to conduct IM operations included an EFB and a CGD, shown in Figure 5.

Each pilot in the two-person crew had their own EFB and CGD, with the EFB mounted on the outboard panel and the CGD just outboard of the aircraft’s navigation display. Since the EFB is mounted outside of the pilot’s optimal primary field of view (defined by the FAA as within ±15° degrees horizontally of a forward line of sight, and within ±15° degrees vertically of a 15° degree downward line of sight), the CGD was mounted within their optimal primary field of view to display to the flight crew the critical subset of information from the EFB needed to conduct the IM operation.

![Figure 4. Integrated Flight Deck (IFD) full-mission simulator.](image)
F. Flight Crew Procedures

During BASELINE operations, flight crews were expected to comply with published procedures and ATC instructions as they do during current real-world operations. During IM operations, a design goal was for the IM flight crew procedures to mirror current day arrival procedures as much as practical. The IM procedures were divided into two distinct phases: first the flight crew programmed the EFB, and then they flew the arrival using the VNAV Speed Mode based on the IM speed displayed on the CGD.

The first phase, programing the EFB, required the subject flight crew to enter the destination airport, arrival and approach procedure, and descent forecast wind into the EFB at any time during the flight but before the IM clearance information was entered. Once ATC issued the IM clearance, the flight crew was required to enter the information for that particular IM operation (for example, the Target call sign, Target route of flight, assigned spacing goal, etc.).

The second phase, flying the aircraft in VNAV Speed using the IM commanded speeds, occurred after all the data had been entered into the EFB, and all the criteria to conduct an IM operation had been met (Target within ADS-B range and both aircraft on a segment of the arrival procedure where the airspeed is specified). With valid data entry and all the IM criteria met, the onboard spacing software displayed the IM commanded speed on the CGD and EFB for the flight crew to set in the mode control panel airspeed window (triggering the VNAV Speed Mode). If the flight crew assessed the airspeed as feasible, the IM procedures were to 1) notify ATC they were accepting the IM clearance, 2) press the EXECUTE button on the EFB to trigger the IM PAIRED mode in the software, and then 3) set the speed in the speed window of the mode control panel.

III. Results and Discussion

To assess the acceptability of the three types of IM operations, data were collected in the form of acceptability, usability, and workload ratings via electronic post-run and post-experiment questionnaires. Additional results from the IMAC experiment can be found in references 4 and 5. For all post-run questionnaire data, statistical analysis was performed using the Wilcoxon signed rank test, a nonparametric test appropriate for analyzing ordinal data. There were no statistically significant differences between the PF and PM responses, so the data were combined for this report.
A. Acceptability of IM Operations

The flight crew acceptability of IM operations was measured using a 7-point Likert rating scale (‘1’ = “Completely Unacceptable” to ‘7’ = “Completely Acceptable”) and data were collected from each of the subject pilots via electronic post-run questionnaires. For all five operations, the mean acceptability ratings (see Table 2) were statistically significantly greater than ‘5’ (p ≤ 0.006), indicating that flight crews found the IM operations to be acceptable on average.

B. Acceptability of IM Commanded Speed

The operational acceptability of IM speed commands was measured using a 7-point Likert rating scale (‘1’ = “Completely Unacceptable” to ‘7’ = “Completely Acceptable”) via electronic post-run questionnaires (see Table 2). For the CAPTURE and MAINTAIN clearances, the mean acceptability ratings were statistically significantly greater than ‘5’ (p ≤ 0.026). For the CROSS clearance and the MIXED scenarios, the mean acceptability ratings were not statistically significantly greater than ‘5’ (p = 0.062 and p = 0.086, respectively), although the observed median was ‘6.’ This indicates that flight crews found the IM commanded speeds to be operationally acceptable for the CAPTURE and MAINTAIN clearances, but were uncomfortable with some of the commanded speeds during the CROSS clearance and the MIXED scenarios.

| Table 2. Flight Crew Acceptability Ratings from Post-Run Questionnaire |
|-----------------|--------------------|--------------------|--------------------|
| Operation       | N   | Mean | SD   | Mean | SD | Mean | SD |
| BASELINE        | 48  | 6.8  | 0.6  | n/a  | n/a | n/a  | n/a |
| CAPTURE         | 46  | 5.6  | 1.4  | 5.4  | 1.7 | 6.2  | 1.2 |
| CROSS           | 48  | 5.7  | 1.1  | 5.2  | 1.5 | 6.3  | 1.0 |
| MAINTAIN        | 42  | 5.8  | 1.5  | 5.5  | 1.7 | 6.3  | 1.1 |
| MIXED           | 48  | 5.6  | 1.6  | 5.3  | 1.7 | 6.2  | 1.1 |

Nineteen of the 24 (79%) pilots reported in the post-experiment questionnaire that the IM commanded speeds caused unexpected or undesired behavior at some time throughout the experiment. In particular, multiple large speed changes are undesirable, and large speed changes near waypoints with altitude constraints sometimes resulted in the aircraft not meeting those constraints (although this issue was most prominent in the lower fidelity ASTOR simulators).

C. Acceptability of Head Down Time

The acceptability of the amount of head down time required of the PM to input information from the IM clearance(s) into the EFB was measured using a 7-point Likert rating scale (‘1’ = “Completely Unacceptable” to ‘7’ = “Completely Acceptable”) via electronic post-run questionnaires. For all IM procedures, the mean acceptability ratings (see Table 2) were statistically significantly greater than ‘5’ (p < 0.0005). Therefore, pilots found the amount of head down time required of the PM to input information from the IM clearance(s) into the EFB to be acceptable.

D. Workload during IM Operations

The flight crews’ overall workload during IM operations was measured using the Modified Cooper-Harper (MCH) subjective workload rating scale via electronic post-run questionnaires. A rating of ‘3’ on the MCH rating scale was determined to be acceptable a priori, where ‘3’ indicates that the instructed task is fair and/or has mild difficulty, and an acceptable level of operator mental effort is required to attain adequate system performance. For all five operations, the mean MCH workload ratings (see Table 3) were statistically significantly less than ‘3’ (p ≤ 0.010). Therefore, flight crews found the workload level experienced to be acceptable on average.
Table 3. Flight Crew Workload Ratings from Post-Run Questionnaire

<table>
<thead>
<tr>
<th>Operation</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASELINE</td>
<td>48</td>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td>CAPTURE</td>
<td>46</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td>CROSS</td>
<td>48</td>
<td>2.4</td>
<td>1.9</td>
</tr>
<tr>
<td>MAINTAIN</td>
<td>42</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>MIXED</td>
<td>48</td>
<td>2.4</td>
<td>1.9</td>
</tr>
</tbody>
</table>

E. Ratings and Comments About Controller-Pilot Phraseology

The acceptability of the use of voice communications was measured using a 7-point Likert rating scale (‘1’ = “Completely Unacceptable” to ‘7’ = “Completely Acceptable”) via electronic post-run questionnaires. For all types of operations, the mean acceptability ratings were statistically significantly greater than ‘5’ (p < 0.0005), indicating the flight crews found the use of voice communications to receive the IM clearances to be acceptable.

Four of 24 (16.6%) pilots responded on the post-experiment questionnaire that they experienced confusion during the IM clearance instruction between the Target versus Ownship call sign, three (12.5%) reported not hearing or noticing the issue, and the remaining 17 (70.8%) reported they did not experience any confusion throughout the experiment. The acceptability of using the Target’s call sign in the IM clearance was also measured using a 7-point Likert rating scale (‘1’ = “Completely Unacceptable” to ‘7’ = “Completely Acceptable”) via electronic post-run questionnaires. The mean acceptability ratings were statistically significantly greater than ‘5’ (p < 0.0005), indicating the flight crews found the use of the Target call sign in an IM clearance to be acceptable.

Comments from the post-experiment questionnaire included:
- suspending or cancelling the IM operation by the air traffic controller sometimes caused confusion, particularly if the controller did not use the specific phraseology to ‘Suspend IM’;
- the flight crew was not always clear when they had to advise the controller they were conducting IM; and
- slightly more than half of the pilots reported using the phonetic alphabet to issue the Target call sign would help with distinguishing it from the IM aircraft call sign, and provide clarity when the call sign was unusual.

F. Ratings and Comments About the IM Interface and Displays

The intuitiveness of entering IM clearance information into the EFB was measured using a 7-point Likert rating scale (‘1’ = “Completely Unintuitive” to ‘7’ = “Completely Intuitive”) via the post-experiment questionnaires. For all types of operations, the mean intuitiveness ratings were statistically significantly greater than ‘5’ (p < 0.0005), indicating the flight crews found the intuitiveness of entering the IM clearance into the EFB to be acceptable. Shown in Figure 6 are the two primary pages on the EFB with which pilots entered information into the IM software system.
The usefulness of the information displayed on the CGD was also measured using a 7-point Likert rating scale (‘1’ = “Not useful at all” to ‘7’ = “Very useful”) via the post-experiment questionnaires. For all types of operations, the mean CGD display usefulness ratings were statistically significantly greater than ‘5’ ($p < 0.0005$), indicating the flight crews found the usefulness of the CGD display to be acceptable. Figure 7 illustrates three examples of the CGD the flight crew rated. In particular, the use of reverse video to highlight a new IM speed (right panel) was noted as useful.

Comments from the post-experiment questionnaire included:

- incorporate aural alerts for speed changes to reduce head down time;
- the FAST/SLOW indicator contained useful information, however the display itself was not intuitive;
- the meaning and operation of EARLY/LATE indicator was not clear to many of the subjects (see Figure 5);
- when there was a delay between when the flight crew entered the IM clearance (the software was armed) and when the operation actually began or commenced (all the criteria had been met), some flight crew reported being unsure if the procedure was operating properly; and
- the lack of cockpit displays indicating the IM spacing operation was no longer feasible was identified as an issue to be addressed prior to implementing the IM procedures in real-world operations.
V. Conclusion

Interval Management Alternative Clearances (IMAC) was a human-in-the-loop simulation experiment conducted to explore the efficacy and acceptability of three IM operations: CAPTURE, CROSS, and MAINTAIN. Two weeks of data collection were conducted, with a total of 24 subject pilots and 8 subject controllers flying a total of 229 arrival operations into Denver International Airport (of which 113 were IM operations).

Overall, the IM operations and procedures were rated very favorably by the flight crew in terms of acceptability, workload, and pilot Head Down time. In general, the flight crew also rated the IM commanded speed as operationally acceptable during the CAPTURE and MAINTAIN operations; however, a few flight crews were uncomfortable with some of the commanded speeds during the CROSS and MIXED operations.

The flight crews also reported that the controller-pilot phraseology developed for the ATD-1 concept of operations was very acceptable. Of particular interest to the research team was the use of the Target’s call sign in the IM clearance, and whether it would cause confusion with either the IM or Target flight crews. Over 70% of the flight crews reported it was not an issue, and they reported that using the phonetic alphabet to issue the Target call sign was an acceptable technique to clearly differentiate the Target call sign from the Ownship call sign.

However, the current instantiation of IM operations in busy arrival operations that use voice to communicate between controllers and flight crew, has several key issues that should be resolved. These issues are:

- the higher frequency of speed changes in an IM operation relative to non-IM operations;
- changes to the commanded airspeed (specifically speed decreases) when approaching an altitude constrained waypoint significantly increased the likelihood the aircraft would not meet the altitude constraint;
- the FAST/SLOW and EARLY/LATE displays were problematic for some flight crews;
- long delays between entering the IM clearance and the IM operation commencing caused some pilots to be unsure if the IM operation was progressing as intended;
- the need for unambiguous cockpit displays indicating the IM spacing operation was no longer feasible.

In summary, the IM spacing algorithm and procedures met the ATD-1 established success criteria and typically did interact as desired with non-IM operations; however, the challenges listed above should be resolved prior to implementing IM into real-world operations.

Acknowledgments

The authors gratefully acknowledge the very hard work and significant amount of time invested by the software development teams at NASA Langley, and the important contributions by the software team at NASA Ames, without either of which this experiment would not have occurred.

References


American Institute of Aeronautics and Astronautics