Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)

Automated Air Traffic Control Operations with Weather and Time-Constraints

A First Look at (Simulated) Far-Term Control Room Operations

Thomas Prevot
NASA Ames Research Center
Moffett Field, CA, USA
Thomas.Prevot@nasa.gov

Jeffrey R. Homola, Lynne H. Martin, Joey S. Mercer and Christopher C. Cabrall
San Jose State University
NASA Ames Research Center
Moffett Field, CA, USA

***Abstract*— In this paper we discuss results from a recent high fidelity simulation of air traffic control operations with automated separation assurance in the presence of weather and time-constraints. We report findings from a human-in-the-loop study conducted in the Airspace Operations Laboratory (AOL) at the NASA Ames Research Center. During four afternoons in early 2010, fifteen active and recently retired air traffic controllers and supervisors controlled high levels of traffic in a highly automated environment during three-hour long scenarios. For each scenario, twelve air traffic controllers operated eight sector positions in two air traffic control areas and were supervised by three front line managers. Controllers worked one-hour shifts, were relieved by other controllers, took a 30-minute break, and worked another one-hour shift. On average, twice today's traffic density was simulated with more than 2200 aircraft per traffic scenario. The scenarios were designed to create peaks and valleys in traffic density, growing and decaying convective weather areas, and expose controllers to heavy and light metering conditions. This design enabled an initial look at a broad spectrum of workload, challenge, boredom, and fatigue in an otherwise uncharted territory of future operations. In this paper, we report human-systems integration aspects, safety and efficiency results as well as airspace throughput, workload, and operational acceptability. We conclude that, with further refinements, air traffic control operations with ground-based automated separation assurance can routinely provide currently unachievable levels of traffic throughput in the en route airspace.**

Keywords- separation, trajectories, automation, NextGen, workload, human-systems integration

I. INTRODUCTION

In this paper we discuss results from a recent high fidelity simulation of air traffic control (ATC) operations with automated separation assurance in the presence of weather and time-constraints. The primary purpose of automating separation assurance is to enable air traffic controllers to manage much higher traffic densities than today. By eliminating airspace capacity constraints resulting from controller workload limitations, automation for separation management can reduce the need for costly traffic management initiatives. Today, whenever air traffic demand exceeds capacity, traffic management initiatives are put in place to reduce the number of aircraft entering congested sectors. In many cases demand is reduced by holding aircraft at their departure airports. These ground stops avoid burning extra fuel and polluting the environment unnecessarily. However, ground delay programs often have a severe impact on airline schedules and inconvenience many passengers. When delays are taken in flight, the aircraft fly longer routes than necessary, which increases the cost and the environmental impact of each flight. The weather impact on airspace throughput often ripples through the National Airspace System (NAS) and results in inefficiencies, long delays, and increased cost.

New approaches to separation management can help alleviate some of these problems. Increasing airspace capacity by automating separation assurance has been studied in some detail over the past decade. Ground-based and airborne concepts involve new automation capabilities and new procedures for the human participants, either controllers or pilots. The primary difference between ground-based and airborne concepts lies in the location of these changes: in ground-based ATC facilities in the first concept, and distributed among aircraft in the other. In the concept *ground-based automated separation assurance* [1][2] ('ground-based concept'), ground-based automation and air traffic controllers manage the separation between all aircraft within a defined airspace. In the concept *airborne trajectory management with self-separation* [3], the pilot manages the separation for his or her aircraft supported by an onboard Airborne Separation Assistance System (ASAS). Detailed descriptions and comparisons of both concepts can be found in [4] and its associated references. In this paper, we will only discuss the ground-based approach studied in this simulation.

In the ground-based concept, air traffic control automation supports and enables the controller to manage more aircraft within the same airspace than today by having the automation -not the air traffic controller- monitor traffic for potential conflicts. Additionally, the automation conducts many workload-intensive routine tasks such as transferring ownership and communication frequencies between air traffic control sectors. Relieved of these tasks, controllers can concentrate on managing the non-routine operations that often require human intelligence, ingenuity, and experience. As a result, more aircraft can be controlled within a given airspace. Airspace saturation occurs at higher traffic levels than today resulting in

This work was sponsored by the NASA Airspace Systems Program's Concepts and Technology Development Project.

fewer aircraft reroutes and ground-stops. More aircraft get their most efficient, user-preferred 'green' trajectories. Passengers experience less delay on busy travel and/or bad weather days.

The paper is organized as follows: First, we describe the operational concept, and then state the research questions of interest for the current study. Next, we describe the method, present the results and discuss key findings. Finally, we outline future work and state our main conclusions.

II. Operational Concept

Ground-based automated separation assurance is a concept that involves a centralized system with ground-side automation components that monitor and/or manage nominal trajectory-based operations of equipped aircraft, while the controller handles off-nominal operations, provides additional services, and makes decisions when human involvement is needed [2]. The separation responsibility resides with the Air Navigation Service Provider (ANSP), here meaning both the air traffic controller and the ground-based automation. The primary difference to today's system is that the ground-based automation is responsible for conflict detection, and separation assurance automation generates and sends conflict resolution trajectories automatically via data link to the aircraft. The controller is involved in routine conflict resolutions only when the automatic trajectory change would impose excessive delay or a drastic altitude change. The flight crews' responsibilities related to separation assurance do not change from current operations.

A. *Enabling Environment*

The concept of automated separation assurance is enabled by integrating controller workstations, ground-based automation, data link, Flight Management System (FMS) automation and flight deck interfaces. The ground automation creates, maintains, and communicates trajectories for each flight. The air traffic environment is generally in line with the mid- to far-term environment for the en route airspace outlined by the Federal Aviation Administration (FAA) [5]. The following characteristics are assumed: each aircraft entering the airspace is equipped with an FMS that meets a required navigation performance (RNP) value of 1.0 and has integrated data link for route modifications, frequency changes, cruise altitudes, and climb, cruise, and descent speeds similar to current-day Future Air Navigation System (FANS) technology. Data link is the primary means of communication, and all aircraft are cleared to proceed, climb, cruise and descend via their nominal or uplinked trajectories. High accuracy surveillance information for position and speed is provided via Automatic Dependent Surveillance Broadcast (ADS-B) or a comparable source. In order to reduce trajectory uncertainties, FMS values for climb, cruise, and descent speeds, as well as weight, are communicated to the ATC system. The goal is to make conflict detection highly reliable and to detect trajectory-based conflicts with sufficient time before any predicted initial Loss Of Separation (LOS). However, some sources of trajectory uncertainties remain, including flight technical differences, trajectory mismatches between the air and the ground, inaccurate performance estimates, and inaccurate weather forecasts used by the air and the ground automation. A conformance monitoring function detects off-trajectory operations and triggers an off-trajectory conflict probe. The trajectory generation function used for conflict resolution and all trajectory planning provides FMS compatible and loadable trajectories. These trajectories account for nominal transmission and execution delays associated with data link messaging. Automated trajectory-based conflict resolutions are generated for conflicts with more than three minutes to initial LOS. When conflicts are detected with less time before LOS, an automated tactical conflict avoidance function generates heading changes and sends them to the flight deck via a separate high-priority data link connection (e.g. Mode-S).

B. *Roles and Responsibilities*

The ANSP is responsible for maintaining safe separation between aircraft. The ground automation is responsible for detecting 'strategic' medium-term conflicts (typically up to 15 minutes) between all trajectories and for monitoring the compliance status of all aircraft relative to their reference trajectory. The ground automation is also responsible for detecting 'tactical' short-term conflicts (typically less than 3 minutes) between all aircraft. The automation sends conflict resolutions automatically via data link to the aircraft whenever predefined tolerances on delay, lateral path, and altitude change are not exceeded. Whenever the ground automation cannot resolve a conflict without controller involvement, it must alert the controller with enough time to make an informed decision and keep the aircraft safely separated. Likewise, the ground automation is also responsible for alerting controllers to other problems and exceptional situations.

Controllers supervise the automation and are responsible for making decisions on all situations that the automation, flight crews or other ANSP operators (i.e., other controllers or traffic managers) present to them. Additionally, they provide service in time-based metering and weather avoidance operations. Issuing control instructions to non data-link-equipped aircraft is also the responsibility of the controller. The controller can use conflict detection and resolution automation to generate new trajectories for any aircraft. Controllers use data link to communicate with equipped aircraft and voice to communicate with non-data-link-equipped aircraft.

Flight crews are responsible for following their uplinked (or initially preferred) trajectory within defined tolerances and for the safe conduct of their flight (like today). Flight crews can downlink trajectory-change requests at any time. The ground automation probes requested trajectories for conflicts without involving the controller. If the requested trajectory is conflict free, the automation uplinks an approval message. Otherwise, it alerts the controller that there is a trajectory request to be reviewed.

C. *Air Traffic Controller Workstation*

Fig. 1 depicts the air traffic controller workstation prototype designed for the above distribution of roles and responsibilities and used for the current study. Aircraft that were managed by the automation within the controller's sector had a brighter icon than the aircraft outside that area, which were dimmed. Additional information in data tags and colors were used to draw the controller's attention to a specific problem. The display was designed for general situation awareness and management by exception. The sector displayed in Fig. 1 contained approximately three times as many aircraft as can be controlled within this sector in current-day operations. All functions for conflict detection and resolution, trajectory planning, and routine operations were directly accessible from the controller display. Transfer of control and communication between sectors was conducted by the automation. Nominally, aircraft were displayed as chevrons with altitudes, a design originally developed for cockpit displays of traffic information [6]. Traffic conflict information, hazard penetration, and metering information was presented where applicable. Full data tags were only displayed in short-term conflict situations, or when the controller selected them manually. Time-based metering was supported via timelines and meter lists. The timelines showed aircrafts' estimated and scheduled arrival times at specific fixes, usually meter fixes into congested airports.

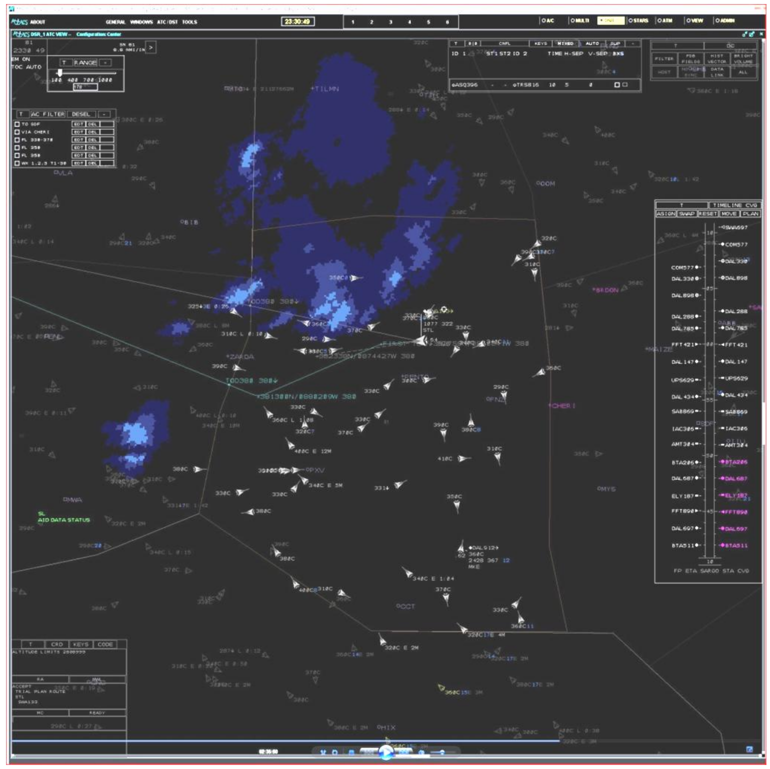

Figure 1. Controller display designed for automated separation assurance.

The controller could request trajectories to avoid traffic conflicts and weather hazards and to meet time-constraints via various easy-to-use mechanisms using keyboard entries, data tag items, the conflict list, or the timeline. The automated trajectory-based conflict resolutions were generated by an autoresolver module originally developed as part of the Advanced Airspace Concept [7]. When initiated by the controller, the automatically generated trajectory became a provisional trial-plan trajectory (e.g., the cyan line in Fig. 1). The controller could then modify and/or uplink the trajectory constraints to the aircraft. The automation immediately probed all trajectory changes for conflicts and the tools provided real-time conflict feedback when used interactively.

When a conflict was predicted to occur within less than three minutes to LOS, the Tactical Separation Assisted Flight Environment (TSAFE) [8] function was activated, which computed heading changes for one or both of the aircraft involved in the conflict. In the current study, the automation automatically sent the heading change(s) at two minutes to predicted LOS. This heading change solved the immediate conflict, but left the aircraft in 'free track' with no trajectory to the destination, requiring the controller to use the trial planning tools to create and send a new trajectory to the aircraft.

III. PROBLEM

Prototype technologies for ground-based automated separation assurance have been developed and studied in fast-time and real-time simulations as well as in laboratory analyses of real air traffic data feeds [9][10][11]. The results to date indicate that building these technologies for operational use is challenging but achievable. Part-task studies of controllers and pilots interacting with existing displays and controls as well as prototypes of future systems have also shown promising results towards developing usable and useful operator interfaces [2]. However, to our knowledge there has been no attempt to investigate the impact and effectiveness of highly automated air traffic control as a routine operating mode in the air traffic control room. Little to nothing is known about whether these operations can create a safe and acceptable work environment for air traffic controllers and front line managers. How do controllers coordinate? How do they change shifts? What information do they need to communicate to each other? It is also not known whether the approach can be effective when there are frequently twice as many aircraft as today in the airspace. What if this airspace is impacted by rapidly changing weather conditions? What if many aircraft have to be transitioned into busy airports?

The purpose of the research described here was to get a first look at simulated far-term control room operations with automated air traffic control, in the presence of weather and time-constraints. The study was designed to provide early insights and initial answers to some of the questions posed above. We summarize the primary research question as follows: *Can air traffic control operations with ground-based automated separation assurance be an effective and acceptable means to routinely provide high traffic throughput in the en route airspace?*

IV. METHOD

The method was to run a high fidelity human-in-the-loop simulation of air traffic control operations with ground-based automated separation assurance. During the simulation, traffic at much higher densities than today transitioned the airspace and had to be sequenced into various nearby airports. The operations were sustained for multiple hours and impacted by convective weather cells that grew, decayed and moved. This long run duration with realistic weather scenarios was chosen to observe operational aspects that are not represented in typical shorter simulations, such as shift changes, stress, boredom, and fatigue. Descriptions of the experiment design, airspace, apparatus, participants, and procedures follow.

A. *Experiment Design*

The experiment was designed as an exploratory study rather than a formal evaluation. Controllers operated in a comprehensive work environment that required them to perform a wide range of air traffic control tasks. Three parameters were varied: (1) traffic demand on the airspace, (2) traffic demand on the metering fixes, and (3) convective weather situation.

The operator stations, tools, and function allocation stayed constant throughout all runs. Traffic demand on airspace and metering fixes was varied within and between runs, with two basic traffic scenarios: (1) a Light Metering scenario with 2216 aircraft, moderate arrival flows with little meter delay and (2) a Heavy Metering scenario with 3060 aircraft, dense arrival flows often requiring more than five minutes of meter delays to be absorbed. Two different weather scenarios were used, where the convective weather was growing or decaying within half of each scenario and absent during the other half. This resulted in four different challenging traffic, weather and metering problems designed to stimulate a wide range of controller activities related to air traffic control and coordination. Each scenario lasted for three hours and for analysis purposes can be divided into three consecutive one-hour long phases. Each phase was a combination of a light or heavy metering situation and the presence or absence of growing or decaying weather. Table 1 shows the design and run schedule.

TABLE 1. EXPERIMENT DESIGN AND SCHEDULE.

	Day1		Day2		Day3		Day4	
	traffic	Wea-ther	traffic	Wea-ther	traffic	Wea-ther	traffic	Wea-ther
Phase 1 1:00 PM	Light	0%	Heavy	2-5 sect. 27%-0 decaying	Light	2-5 sect. 27%-0% decaying	Heavy	0%
Phase 2 2:00 PM	Mete-	0-3 sect. 0- 17% growing	Mete-	0-2 sect. 8%-0% decaying	Mete-	0-2 sect. 8%-0% decaying	Mete-	0-3 sect. 0-17% growing
Phase 3 3:00 PM	ring	3-5 sect. 0-27% growing	ring	0%	ring	0%	ring	3-5 sect 0-27% growing

For example, day 1 was a Light Metering day with weather starting to grow in Phase 2, impacting three sectors with the most impacted sector reaching 17% weather coverage. In Phase 3, weather kept growing, impacted five sectors and covered up to 27% of the airspace in the sector impacted most.

B. *Airspace*

The simulation was situated in the central United States and covered eight high altitude sectors: four on the eastern side of Kansas City Center (ZKC) and four on the western side of Indianapolis Center (ZID), as shown in Fig. 2. To create challenging metering problems, arrivals into various airports were scheduled over certain meter fixes such that they could conduct optimum profile descents from the en route airspace. Airports with meter fix time-constraints included BNA, CVG, MSP, ORD, SDF, and STL.

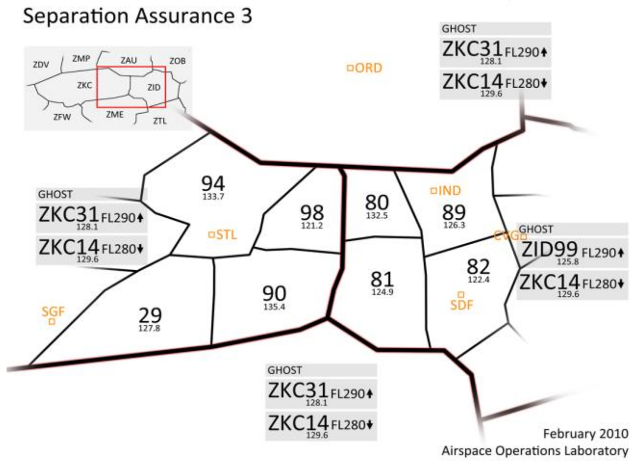

Figure 2. Test Airspace.

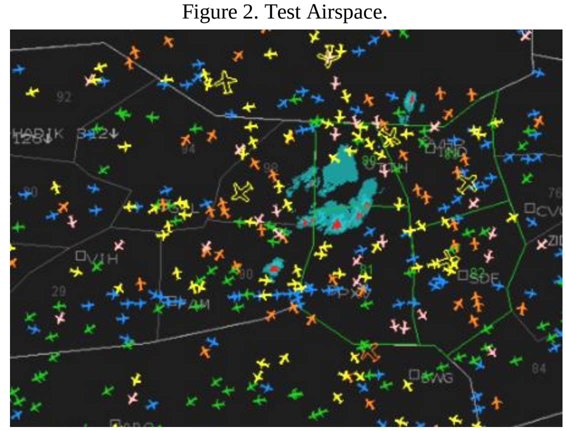

Figure 3. Scene from simulation (Light Metering, growing weather).

Fig. 3 shows a scene as it was displayed on an overhead projector in the ZID control room during Phase 3 on Day 1 of the simulation. Weather impacted four sectors, two of them severely, forcing controllers to route the traffic around the weather cells. In this situation the weather coverage of ZID 80 was about 20%, which made about half the sector unusable.

C. *Apparatus*

The simulation was conducted in the Airspace Operations Laboratory (AOL) at the NASA Ames Research Center [11]. The AOL's Multi Aircraft Control System (MACS) software was used for all simulation and rapid prototyping activities [12]. MACS provides high-fidelity display emulations for air traffic controllers and managers as well as user interfaces and displays for confederate pilots and flight crew participants, experiment managers, analysts, and observers. Scenario and target generation capabilities are also built into MACS, which were used to generate and run the traffic and weather problems. MACS' integrated data collection system was used to collect the quantitative measures of interest at each operator station as well as overall traffic progression, including aircraft states, conflicts, and sector counts.

In order to provide the required automation support to the controller, a new NextGen ATC workstation prototype was developed based on an emulation of the operational en route controller system. The workstation provided access to key functions that supported the operator in managing high traffic densities effectively. Fig. 1 earlier in this paper shows the

Figure 4. Air traffic control room in the AOL.

controller display as implemented in MACS and used for this research.

For this study, the AOL was configured with two participant control rooms, each hosting the four air traffic control sector positions and one supervisor position in ZID and ZKC, respectively. Fig. 4 shows one of the air traffic control rooms with four radar positions and the supervisor workstation. Each workstation displayed one sector that was worked by a single radar (R-Side) controller.

D. *Participants*

Six active FAA front line managers that were certified as current on the radar position were complemented by six recently retired air traffic controllers and one supervisor from Oakland Center. Together, they staffed the eight air traffic control and two area supervisor positions in the two air traffic control rooms. Three additional confederate controllers worked the traffic flows into and out of the test sectors, and ten general aviation pilots served as pseudo pilots, who operated the simulated traffic.

E. *Experimental Procedure*

After three days of training, data were collected during the afternoons on four consecutive days, when a three-hour long scenario with either 2,216 or 3,060 aircraft was run. In each run, four teams of three controllers rotated through two neighboring sectors, so that each controller worked each sector for one hour. The rotation was scheduled such that a controller had a 30 minute break after each shift and was therefore never on position for longer than one hour. Shift changes were scheduled and posted in the control room and the break room. During each shift change, the outgoing controller briefed the incoming controller, who then signed into the workstation.

System data as well as user inputs were recorded with the MACS data collection system. At three-minute intervals throughout each run, participants were prompted visually and audibly to rate their perceived workload. The position-relief briefings were recorded with the voice communication system. The sign-in/sign-out process at the shift change recorded the exact time at which a new operator took over a position. At the end of shifts in the first two phases of each afternoon the outgoing participants responded to a short questionnaire in the break room. After Phase 3, all participants completed a more comprehensive post-run questionnaire that included items on function allocation. All questionnaires (post-shift, post-run, and post-simulation) were posted electronically.

V. RESULTS

In this section, we present results on airspace capacity and throughput, controller workload, safety, efficiency, acceptability, and function allocation.

A. *Airspace Capacity and Throughput*

Table 2 presents the mean aircraft count per sector within the eight-sector test airspace for the three phases of each run accompanied by the standard deviations. The results show that the mean number of aircraft in each sector was much higher than is experienced today, particularly for the counts in Phases 2 and 3. Table 2 also shows that the weather had little impact on the aircraft count, indicating that high throughput was maintained in the presence of weather. The peak aircraft count in the peak sector within the test area provides a more striking depiction of the elevated traffic levels that were experienced and managed by the participants.

TABLE 2. MEAN AIRCRAFT COUNT PER TEST SECTOR FOR THE LIGHT AND HEAVY METERING CONDITIONS.

	Phase 1			Phase 2			Phase 3		
	AC Count	*SD*	Peak	AC Count	*SD*	Peak	AC Count	*SD*	Peak
Light Meter.-Decaying Wx	**19.0**	*2.7*	32	**27.5**	*5.9*	48	**25.9**	*5.0*	42
Light Meter.-Growing Wx	**19.4**	*3.8*	36	**27.1**	*5.7*	51	**21.2**	*6.5*	44
Heavy Meter.-Decaying Wx	**20.1**	*7.4*	52	**25.3**	*8.7*	50	**29.1**	*7.4*	60
Heavy Meter.-Growing Wx	**19.6**	*8.0*	48	**24.1**	*8.1*	47	**28.1**	*8.6*	62

Fig. 5 presents time-series plots of the peak aircraft counts in the peak sectors within the ZKC and ZID test areas throughout the three-hour runtime in the Heavy Metering condition that show the detailed characteristics of the traffic load. It shows that in the ZKC area, there were sectors that experienced aircraft counts between 40 and 50 for sustained periods of time, and one sector in particular in ZID experienced counts above 60 aircraft. As a reference point, today the peak aircraft count for these sectors is not supposed to exceed the Monitor Alert Parameter (MAP) of 18 aircraft.

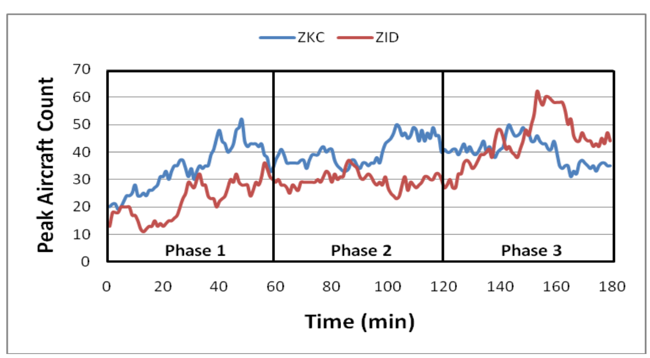

Figure 5. Peak aircraft counts for the ZKC and ZID test areas in the Heavy Metering condition.

B. Controller Workload

1) Real-time Ratings

The real-time workload ratings were on an interval scale from one to six, with six representing the highest level of workload possible. Fig. 6 presents the overall mean workload reported by the R-side test participants in each of the Metering-Weather conditions across the three phases of each run.

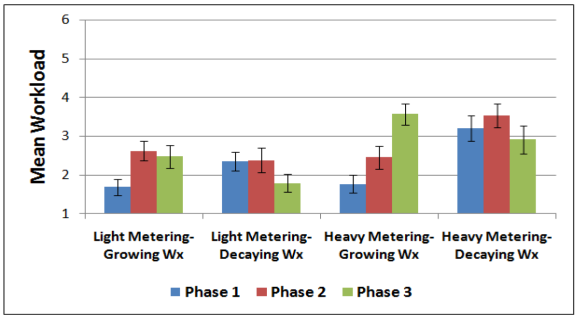

Figure 6. Mean reported workload for each Metering-Weather condition and run phase.

From these results and the post-run ratings discussed in the next section (see Fig. 8), it appears that the workload increased with more severe weather and metering conditions. This is not surprising given that there were controller's tasks associated with the aircraft that required metering and weather reroutes. In contrast, the raw aircraft count does not appear to correlate with workload. This is indicated in Fig. 6 where a phase without convective weather received consistently the lowest mean workload rating in each run independent of aircraft count. Additional evidence is given in Fig. 7, which presents the mean workload reported by the ZKC R-sides overlaid on the mean AC counts for the Heavy Metering runs. Phase 1 of the Growing Weather run (upper portion) did not involve any weather cells, and the mean workload was relatively low despite high levels of traffic. In contrast, the workload reported for Phase 1 in the Decaying Weather run (lower portion) was much higher despite nearly identical AC counts. The only difference was that Phase 1 of the Decaying Weather run

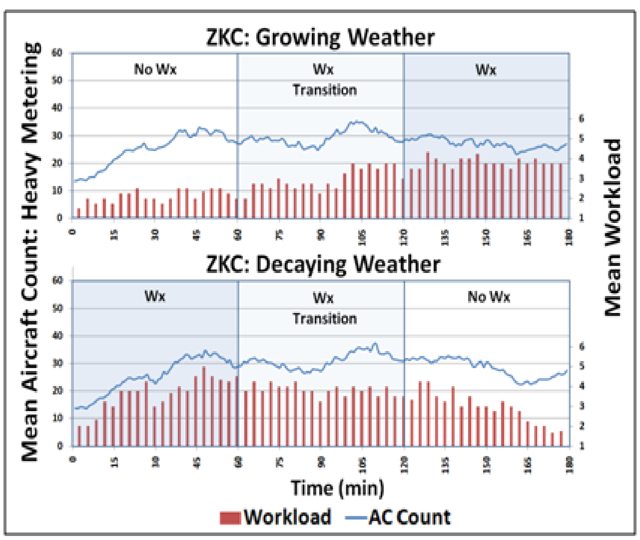

Figure 7. Mean workload overlaid with mean AC count in the Heavy-Growing and Decaying run for the ZKC test area.

started with weather cells affecting the test airspace whereas weather affected later phases in the Growing Weather run.

2) Post-Run Workload Ratings: NASA TLX

In addition to runtime workload ratings, participants provided assessments of their workload following each phase and at the conclusion of each run. Participants completed two of six workload ratings – mental load and time pressure – that form the NASA-TLX workload scale [13] after Phases 1 and 2 in each run. They completed the full TLX scale after the third phase. In each case, the scale ran from 1 (very low) to 7 (very high).

Comparing mental workload and time pressure by the metering and weather conditions showed that on average the Heavy Metering condition (whichever phase it occurred in) was always rated as producing a higher workload than the Light Metering condition. When there was weather, workload was rated as higher than when there was none. A Friedman test showed significant differences between participants' responses on both post-run scales for mental workload ($\chi^2(3)=12.87$, $p=.005$) and time pressure ($\chi^2(3)=13.79$, $p=.003$).

Fig. 8 illustrates the mental workload mean rating for the four conditions (the graph for the time pressure variable is similar). When the weather and metering variables were tested separately using a Wilcoxon Signed Ranks Test, both the presence of weather and the heavy metering significantly increased participants' mental workload ratings (weather: $M_{no\text{-}weather} = 3.67$, $M_{weather} = 4.75$, $Z=3.27$, $p=.001$), (metering: $M_{heavy\text{-}metering} = 4.62$, $M_{light\text{-}metering} = 3.41$, $Z=3.38$ $p=.001$), supporting the real-time workload findings. However, although the level of metering was related to a significant difference in participants' time pressure responses ($p=.000$; $M_{heavy\text{-}metering} = 3.67$, $M_{light\text{-}metering} = 2.12$), the presence of weather was not.

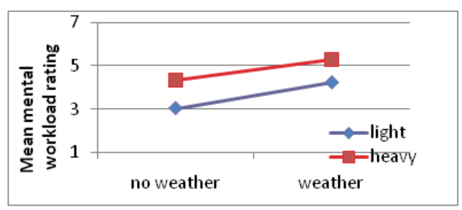

Figure 8. Mean mental workload across the four study conditions.

Participants' general comments on the questionnaires indicated that workload varied considerably depending upon the weather and metering conditions. After phases with Light Metering and no weather, participants said the run was "dull and boring" and they "never had to step in" to assist the automation; after phases with weather and Heavy Metering, participants said "the workload was pretty intense" and that runs were "very busy due to weather reroutes".

C. Safety

1) Losses of Separation

A LOS was recorded anytime two aircraft were simultaneously closer than 5 nmi laterally and less than 800 feet apart vertically. To be included in the following analysis a LOS had to occur within the tests sectors after the first 5 minutes of a run and last for at least 12 consecutive seconds.

These LOS events were further categorized into Operational Errors (OE) and Proximity Events (PE) based upon the lateral separation at the closest point of approach (CPA) measured between the aircraft. If that distance was between 4.5 nmi and 5.0 nmi, the LOS was counted as a PE; whereas if that distance was less than 4.5 nmi the LOS was counted as an OE.

Across the 12 hours of simulation, a total of 1450 LOS events were scripted to occur inside the test airspace, 325 in each Light Metering Scenario and 400 in each Heavy Metering Scenario. 42 LOS events actually occurred. Of these, 8 were PE and 34 were OE. Fig. 9 shows the number and kind of LOS per weather/metering condition. Initial examinations including video-based analyses were undertaken to broadly characterize LOS in terms of sector counts, weather, phase, shift changes, altitude geometries, locations, cause and severity.

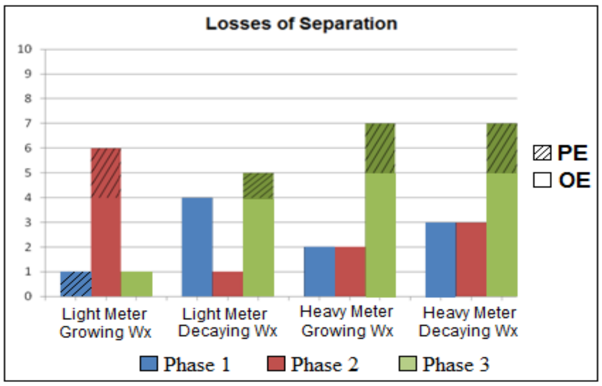

Figure 9. LOS for each phase and weather/metering condition.

Neither the aircraft count nor the amount of weather present within a sector at the time of a LOS appeared to affect the probability of a LOS occurrence. The sector aircraft counts for the 10 minutes prior to a LOS were averaged for each LOS, and this distribution of pre-LOS sector aircraft counts (Min=9.3, Max=43.2, M=26.9, SD=8.7) was seen to be generally representative of the full set of sector aircraft counts seen across all runs (Min=4, Max=62, M=23.9, SD=9.1). Weather was present in the sector of the LOS 11 times, but only five of these occurrences involved a situation where 10% or more of the sector was covered by weather in the minutes leading up to and during the LOS.

Interestingly, it initially appears that time factors might have contributed to LOS events. Regardless of the specific run condition, the majority occurred in Phase 3 (20 LOS events) compared to Phase 2 (12) and Phase 1 (10). Additionally, with respect to the controller rotation, 31% took place within either the first 10 minutes or last 3 minutes of a controller's shift.

Locations and altitude geometries revealed a significant impact of arrival/departure flows of aircraft on the occurrence of LOS events in the simulation. A clear majority (62%) of LOS events were located within portions of ZKC98 and ZID81 with traffic going to/from the STL and SDF airports respectively. Both aircraft were level at cruise altitude in only nine LOS events; all others involved at least one aircraft that was descending (25) or climbing (8). This supports the common understanding that transitioning aircraft pose the biggest challenge to current conflict detection/resolution algorithms. Video recordings of the radar scopes (as well as radio communications) for each LOS were reviewed to assess potential causes. Allowing for a LOS event to have more than one cause, causes were initially attributed as follows: pseudo pilot mistakes (5%), controller judgment/error (12%), conflicting resolution overlap between controller and automation (12%), insufficiencies in trajectory-based conflict resolutions/trial-planning (19%), secondary conflict interactions during off-trajectory operations resulting from prior tactical conflict resolutions (i.e. TSAFE) (24%), and last-minute or no conflict detection (64%).

Fig. 10 illustrates the relationship between cause and closest point of approach (CPA) of the LOS events. The results indicate that the majority of less severe LOS events (CPA > 3 nmi) are caused by conflict detection problems that can likely be avoided by expanding the buffers around the separation minimum of 5 nmi when probing for conflicts. Conflict detection problems, off-trajectory operations following tactical TSAFE resolutions, and controller judgment errors contributed to the most severe LOS events. This result indicates that additional research emphasis needs to be placed on the controller/automation function allocation in short-term conflict situations.

2) *Weather Penetration*

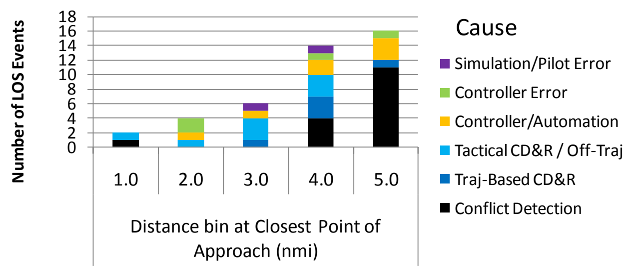

Figure 10. Number of LOS events per distance bin of CPA and cause for LOS (e.g. a LOS with a CPA of 2.3 nmi appears in the 3.0 distance bin).

Instances of aircraft penetrating convective weather provide another safety measure. As mentioned, there were two types of weather patterns used in both the Light and Heavy Metering conditions: Decaying and Growing. The decaying weather pattern was present both in and near the test airspace at start time and gradually dissipated over the course of the first 90 minutes. The growing weather pattern appeared as a smaller collection of cells at the 90th minute and amassed over the final 90 minutes of a run. These patterns were composed of three intensity levels (low, medium, and high), differentiated on the controller displays by color. Throughout each of the runs, the participants were asked to use lateral reroutes to avoid the weather. The controllers used a 'time to weather penetration' indication in the aircrafts data tags to assess when an aircraft needed to be rerouted, and interactive trajectory automation to plan the weather reroute. Both based their weather prediction upon an imperfect weather forecast model that predicted that the current weather moved linearly without changing its shape, while the actual weather changed its shape and direction every six minutes. Therefore reroutes that initially appeared clear of weather, could lead to a weather penetration a few minutes later, because the weather behaved differently than predicted by the linear forecast model.

The number of minutes that an aircraft was in weather at any intensity level was used as the measure of comparison for the weather penetration analysis and is referred to as penetration minutes. Table 3 and Fig. 11 describe how many penetration minutes were scripted (green) into each scenario and how many penetration minutes actually occurred (blue) with controllers working the traffic.

TABLE 3: AIRCRAFT MINUTES IN WEATHER BY CONDITION

Weather / Traffic	Decaying		Growing	
	Scripted	Actual	Scripted	Actual
Light Metering	329	45	502	144
Heavy Metering	233	51	664	325

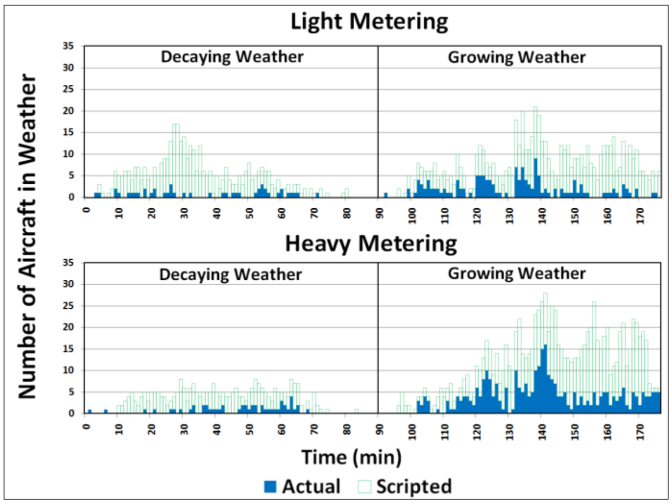

Figure 11. Timeline of weather penetrations per condition.

The totals in Table 3 and the time series plot in Fig. 10 indicate that the controllers were able to avoid weather penetrations almost entirely in the decaying weather problems. The growing weather patterns posed a greater challenge, since the underlying forcast model estimated the size of each weather cell to stay constant while it was actually growing. In the Light Metering condition controllers were still able to reroute all but 29% (144 of 502) of the aircraft succesfully, but the complexity and workload in the Heavy Metering condition caused 49% (325 of 664) of the scripted weather penetrations to actually occur.

D. *Efficiency*

1) *Lateral Path Deviation*

An initial investigation into efficiency was conducted using the amount of lateral path deviation (away from their original flight plan) recorded for each flight. Trajectory changes issued for strategic medium-term conflicts, tactical short-term conflicts, weather avoidance, and for schedule conformance, issued either by the controller or the automation, can all impact lateral path deviation.

When comparing the mean path deviation between the Light and Heavy Metering conditions, the data are similar, with mean values of 0.67 nmi and 0.76 nmi of extra path length, respectively (Fig. 12, left). This suggests that even in the high levels of dense traffic experienced in the Heavy Metering conditions, there was still sufficient maneuverability in the airspace. This finding also indicates that, under this concept of

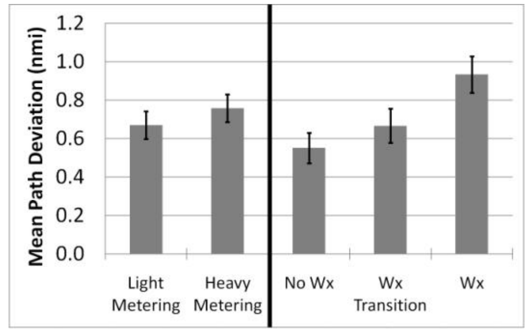

Figure 12. Mean Path Deviation for the two traffic demand levels (left), and the three weather phases (right).

operations, large increases in traffic levels in the NAS can possibly be accommodated without loss of efficiency. The right side of Fig. 12 shows the mean lateral path deviation as a function of weather. Not surprisingly, as more weather is present, path deviations increase, albeit slightly. This may support the real-time workload results; however these insights are preliminary – more detailed analyses of the lateral path deviation are still in progress.

2) *Schedule Conformance*

As the aircraft that were scheduled over meter fixes feeding congested airports left the test airspace, their Estimated Time of Arrival (ETA) was compared to their Scheduled Time of Arrival (STA). This measure was used to determine how well the controllers were able to deliver aircraft according to the arrival schedule. Fig. 13 depicts the schedule conformance. Similar to the path deviation data, there was little difference between the Light and Heavy Metering conditions. On average, scheduled aircraft arrived at their meter fix 7.56 s and 5.97 s later than their STA, respectively. This finding indicates that

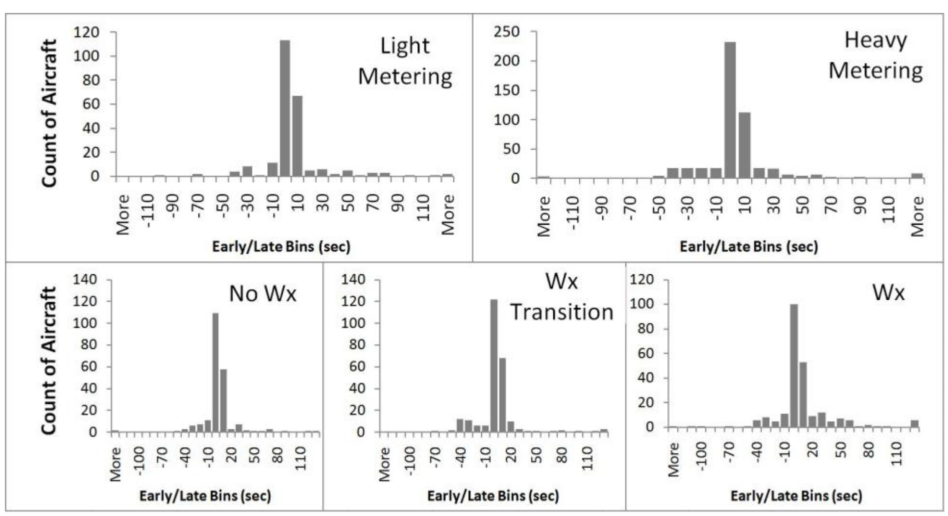

Figure 13. Histograms showing the distribution of ETA-STA discrepancies for the two traffic demand levels (upper) and the three weather phases (lower).

the increase in traffic demand into congested airports and the increase in metering delay to be absorbed did not prevent the controllers from consistently delivering aircraft on schedule. The distribution of this schedule conformance data is presented in the upper histograms in Fig. 13

As to the effect of the presence of weather on schedule conformance, the data show that as more weather is present, the controllers tended to deliver aircraft later relative to the STA. Given that negative values represent an aircraft arriving early at its meter fix, and positive values represent an aircraft arriving late at its meter fix, mean schedule conformances observed were 1.98 s, 7.24 s, and 10.08 s for the No Wx, Wx Transition, and Wx phases, respectively. This is expected, given that multiple metered flows in the scenarios were at some point completely obstructed by the weather cells. The distribution of this data, seen in the lower portion of Fig. 13, is consistent with both the lateral path deviation data and the real-time workload ratings.

E. *Acceptability*

Six of the post-run questions formed an acceptability scale which followed the Controller Acceptance Rating Scale (CARS) developed by [14] as closely as possible. Although the first question was mandatory, the other questions were conditional upon previous answers. Participant answers were compiled to form a scale from one to ten where "1" indicated that the SA operation was not safe through to "10" indicating the operation was acceptable. The CARS ratings were compared over the three phases of each run. On average participants found the SA operations slightly less acceptable as the run progressed, that is, the highest mean acceptability score was reported in Phase 1 (M=7.15, "Moderate compensation required to maintain adequate performance") and the lowest mean CARS was in Phase 3 (M=6.56, "Considerable compensation required to maintain adequate performance"). These differences were statistically significant ($\chi^2(2) = 6.73$, $p=.035$) when tested with a Friedman statistic. This is due to operations being rated as more acceptable in the first phase than in the third. Shown in Fig. 14, there were 12 "uncontrollable" responses in phase 3 (24.9%) versus only 5 in phases 1 and 2 (15.6%).

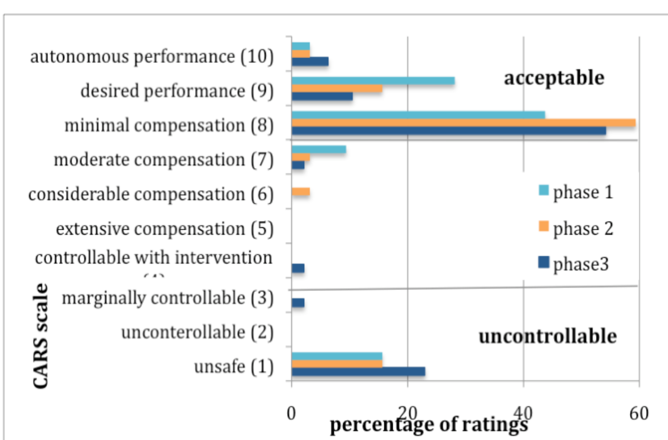

Figure 14. Percentage of CARS ratings in each scale bin for the three phases in each run.

F. *Functional Allocation between Controller and Automation*

In the third phase questionnaire, a question asked participants whether there were tasks that they would have rather done themselves or whether there were tasks that they would have liked the automation to perform. The question about additional tasks that participants would rather perform themselves was asked 33 times. For 17 of these opportunities (51%) participants identified tasks that they would like to do themselves. This suggests that in the other cases (16), although a participant thought s/he had only 'few' or 'some' tasks, s/he did not feel that s/he needed to take control of any more functions. The question about allocating additional tasks to the automation was asked 13 times. In all 13 cases, participants identified functions that they would like to see automated, indicating that they felt they had too much to do. The bar chart in Fig. 15 shows how often a participant voted that a function should be reallocated between themselves and the automation.

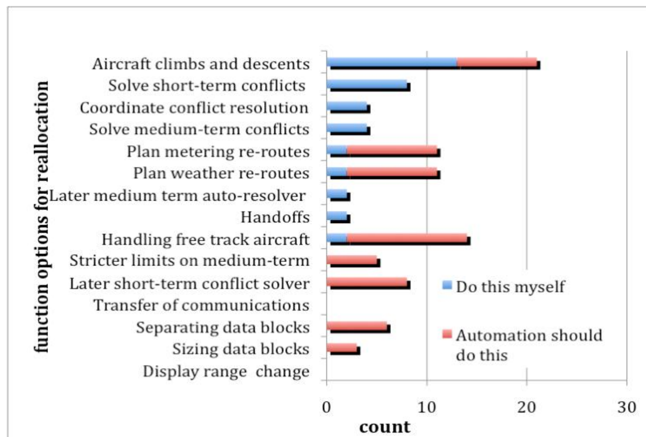

Figure 15: Count of function re-allocation votes after phase 3 in each run.

The most popular function that participants wanted to complete themselves was approval of aircraft climbs and descents (13 of 17= 76%), followed by manually solving short-term conflicts (8 of 17= 47%) and manually solving medium-term conflicts (4 of 17= 23%). No one preferred manual transfer of communications, stricter limits on medium-term conflict solutions, or later auto-solving of short-term conflicts to give them a larger role in these tasks. The most popular candidate for automating was putting free-track aircraft back on a 4D trajectory (12 of 13= 92%), followed by trial planning, and sending weather and metering reroutes (9 of 13= 69%). No participant wanted the range on their display to change automatically and few participants wanted data-blocks to automatically expand or collapse (3 of 13= 23%).

VI. Discussion

The results show that with this concept, sustained high capacity is achievable, even in the presence of convective weather and with heavy metering constraints. They also confirm the results from earlier studies showing that, unlike today, aircraft count is no longer the primary limiting factor and many more aircraft than today can be controlled. However, safety remains an issue, highlighting the importance of robust and reliable automation. While some of the LOS data seem to suggest valid challenges for human operators (e.g., greater numbers in the last phase of any day and also near shift transitions), the majority of LOS events were associated with the problematic complexity that comes with dense departure and arrival flows and fundamentally automation, rather than human, issues. Recognizing this, the proper balance must also be struck between the roles of humans and automation in this concept to maintain a consistent and appropriate level of engagement for the controllers.

This study provided a glimpse of how an air traffic management system works over time, something not seen in most studies because they have runs that are (by necessity) too short in duration. The results show that some events take some time to recover from, and even when a problematic condition no longer exists, its effects can be seen in both controller workload reports and LOS counts. For example, the increase in reported workload matches the onset of weather but also persists after the weather is gone, implying that the controllers needed time to recover from its effects. The acceptability ratings also imply that using a tool for an extended period of time is not the same as using it for a few minutes. There were many more "uncontrollable" ratings in Phase 3 compared to Phase 1 and Phase 2. Testing a tool over time is important because it may highlight aspects of its functionality that need to be attuned for long-term use. Also, the study revealed many human-automation interaction issues, in particular with regard to short-term conflicts, and these require further research.

VII. Future Work

Future work will continue on many levels. The automation in our prototype will be improved to address the conflict detection/resolution deficiencies uncovered in this study, and to provide the additional functionality requested by the controllers. Research, technology, and procedure development will continue to improve the function allocation between air and ground and automation and controllers. Mixed equipage operations and off-nominal situations will also be studied.

VIII. Conclusions

The results from this study show that air traffic control operations with ground-based automated separation assurance can routinely provide currently unachievable levels of traffic throughput in the en route airspace. The ground-based automation system was stress tested in a highly dense and complex environment in the presence of heavy metering constraints and convective weather for sustained periods of time and performed very well overall. Controllers were able to work under this concept of operations in a realistic environment, and found it largely acceptable. Based on the results the automation will be further improved to address safety issues associated with complex traffic situations, and new human-automation integration considerations will inform future work. When implemented properly, these operations can eliminate many airspace capacity constraints and significantly reduce inefficiencies, delays, as well as the environmental impact and cost of air travel.

Acknowledgment

The authors would like to thank Dr. Heinz Erzberger and Dr. Karen Heere for their insights and their support in integrating the autoresolver and TSAFE modules. The experiment was only possible through the excellent work of the entire research, development, and support staff in the AOL. Additional thanks go to NASA's Airspace System Program, the FAA and particularly the controller participants.

References

[1] H. Erzberger, "Transforming the NAS: The next generation air traffic control system", 24th International Congress of the Aeronautical Sciences, Yokohama, Japan, 2005.

[2] Prevot, T., Homola, J., Mercer, J., Mainini, M. and Cabrall, C. "Initial evaluation of NextGen air/ground operations with ground-based automated separation assurance," Eighth USA/Europe Air Traffic Management Research and Development Seminar, Napa, CA, 2009.

[3] D. Wing, "A potentially useful role for airborne separation in 4D-trajectory ATM operations," AIAA 2005-7336, 5th AIAA Aviation Technology, Integration, and Operations Conference (ATIO), Arlington, Virginia, 26-28 September 2005..

[4] D. Wing, T. Prevot, J. Murdoch, et al. "Comparison of ground-based and airborne function allocation concepts for NextGen using human-in-the-loop simulations," AIAA-2010-9293 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Fort Worth, Texas, Sep. 13-15, 2010.

[5] FAA, NAS Enterprise Architecture Portal https://nasea.faa.gov 2010

[6] W. Johnson, V. Battiste, S. Delzell, S. Holland, S. Belcher, and K. Jordan, "Development and demonstration of a prototype free flight CDTI," SAE/AIAA World Aviation Congress, Anaheim, CA, 1997.

[7] H. Erzberger, T. Lauderdale, and Y-C Chu, "Automated conflict resolution, arrival management and weather avoidance for ATM," 27th International Congress of the Aeronautical Sciences, Nice, France, 2010.

[8] H. Erzberger H and K. Heere Algorithm and operational concept for solving short range conflicts. International Congress of the Aeronautical Sciences 2008-8.7.5, Anchorage, Alaska, 2008.

[9] Farley, T. and Erzberger, H. "Fast time air traffic simulation of a conflict resolution algorithm under high air traffic demand," Proc UAS/Europe ATM R&D Seminar, Barcelona, Spain, 2007.

[10] D. McNally and C. Gong, "Concept and laboratory analysis of trajectory-based automation for separation assurance," AIAA 2006-6600, AIAA Guidance, Navigation, and Control Conference and Exhibit, Keystone, Colorado 21 - 24 August 2006.

[11] T. Prevot, P. Lee, T. Callantine, J. Mercer, J. Homola, N. Smith, et al.. "Human-in-the-loop evaluation of NextGen concepts in the Airspace Operations Laboratory," AIAA-2010-7609 AIAA Modeling and Simulation Technologies Conference, Toronto, Ontario, Aug. 2-5, 2010.

[12] T. Prevot., "Exploring the many perspectives of distributed air traffic management: The Multi-Aircraft Control System MACS," In S. Chatty, J. Hansman, & G. Boy. (Eds). HCI-Aero 2002, AIAA Press, Menlo Park, CA. pp 149-154, 2002.

[13] S. Hart, and L. Staveland, "Development of the National Aeronautics and Space Administration (NASA) Task Load Index (TLX): Results of empirical and theoretical research," In P. Hancock & N. Meshkati (Eds), Human mental workload (pp139-183). Amsterdam, North Holland, 1988.

[14] K. Lee, K. Kerns, R. Bone, and M. Nickelson, "Development and validation of the Controller Acceptance Rating Scale (CARS): Results of empirical research," 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, NM, December 2001.

Author Biographies

Dr. Thomas Prevot earned his doctorate in Aerospace Engineering from the Munich University of the German Armed Forces in 1995. **Jeffrey Homola** earned his M.S. in Human Factors and Ergonomics from San Jose State University in 2008. **Dr. Lynne Martin** earned her Ph.D. in Applied Social Psychology at the University of Surrey in 1995**. Joey Mercer** earned his M.S. in Human Factors and Ergonomics from SJSU in 2007. **Christopher Cabrall** earned his M.S. in Human Factors and Ergonomics from SJSU in 2010