

A Human-in-the Loop Exploration of the Dynamic Airspace Configuration Concept

Jeffrey Homola¹, Paul U. Lee², Thomas Prevôt³,
Hwasoo Lee¹, Angela Kessel², and Connie Brasil¹
SJSU / NASA Ames Research Center, Moffett Field, CA, 94035

Nancy Smith⁴
NASA Ames Research Center, Moffett Field, CA, 94035

An exploratory human-in-the-loop study was conducted to better understand the impact of Dynamic Airspace Configuration (DAC) on air traffic controllers. To do so, a range of three progressively more aggressive algorithmic approaches to sectorizations were chosen. Sectorizations from these algorithms were used to test and quantify the range of impact on the controller and traffic. Results show that traffic count was more equitably distributed between the four test sectors and duration of counts over MAP were progressively lower as the magnitude of boundary change increased. However, taskload and workload were also shown to increase with the increase in aggressiveness and acceptability of the boundary changes decreased. Overall, simulated operations of the DAC concept did not appear to compromise safety. Feedback from the participants highlighted the importance of limiting some aspects of boundary changes such as amount of volume gained or lost and the extent of change relative to the initial airspace design.

Nomenclature

<i>ATC</i>	=	Air traffic controller
<i>DAC</i>	=	Dynamic Airspace Configuration
<i>DAU</i>	=	Dynamic Airspace Unit
<i>DFPA</i>	=	Dynamic Fix Posting Area
<i>DSR</i>	=	Display System Replacement
<i>FAA</i>	=	Federal Aviation Administration
<i>FPA</i>	=	Fix Posting Area
<i>IFR</i>	=	Instrument Flight Rules
<i>JPDO</i>	=	Joint Planning and Development Office
<i>MACS</i>	=	Multi-Aircraft Control System
<i>MIP</i>	=	Mixed Integer Programming
<i>NAS</i>	=	National Airspace System
<i>NASA</i>	=	National Aeronautics and Space Administration
<i>NextGen</i>	=	Next Generation Air Transportation System
<i>RNAV</i>	=	Area navigation
<i>RNP</i>	=	Required Navigation Performance
<i>TMU</i>	=	Traffic Management Unit

¹ Research Associate, Human-Systems Integration Division, NASA ARC, MS 262-4.

² Senior Research Associate, Human-Systems Integration Division, NASA ARC, MS 262-4.

³ Senior Research Engineer, Human-Systems Integration Division, NASA ARC, MS 262-4.

⁴ Senior Research Psychologist, Human-Systems Integration Division, NASA ARC, MS 262-4.

I. Introduction

In the National Airspace System (NAS) of today, one of the primary goals and functions of air traffic management resides in allocating and maintaining the balance between air traffic demand and the current and predicted airspace capacity. This is done through a collaborative process involving the various operators among the air traffic organizations (e.g. Command Center, Traffic Management Units (TMU), En route, Terminal Radar Approach Control, and Tower facilities). If the current or forecast demand is expected to exceed capacity due to situations such as convective weather, special events, or controller workload, then a variety of methods are available for use in addressing the demand-capacity imbalance often by reducing the demand to meet the capacity. Some of these methods include enacting traffic management initiatives that involve placing certain aircraft in miles-in-trail or on playbook routes, rerouting aircraft outside of particular sectors or areas, and implementing a ground stop or ground delay program.

While necessary, one common criticism of these traffic management methods is that they are often overarching in their scope and consequently introduce unnecessary and excessive inefficiencies into the system. This problem will likely be compounded if the current FAA predictions for a threefold increase in traffic by the year 2025 hold true (FAA, 2009).

To address these issues and develop viable alternatives, the areas of air traffic controller workload and airspace sector design are often cited as key considerations in any move forward. Because air traffic controller workload currently underlies, in part, the need for and rationale behind implementing traffic management initiatives and is seen as a potential bottleneck jeopardizing the ability for the NAS to accommodate increased levels of traffic, a number of research efforts and concept developments have been aimed at finding ways of reducing air traffic controller workload while being able to provide greater efficiency within the system.

One area of research referred to as the Multi Sector Planner (MSP) examines the role of a position working more tactically than the TMU to an area while simultaneously working more strategically than the area supervisor and the floor (Lee et al., 2006, Corker, Liang, Lee, & Prevôt, 2007, PHARE, 1997). The potential benefit for such a position could be more responsive and dynamic traffic management that would allow for greater efficiencies relative to current management methods as well as reduced impact to system users and a better distribution of workload and resources at the sector level. However, the actions of such a position still require aircraft and flows to be moved in response to demand-capacity imbalances.

An alternative to this approach is instead to look at ways of designing the airspace such that capacity can be increased to meet demand. This relatively recent area of research is referred to as Dynamic Airspace Configuration (DAC), which encompasses three focus areas: restructuring airspace, adaptable airspace, and generic airspace (Kopardekar, Bilimoria, & Sridhar, 2007). With respect to the first area: restructuring airspace, a number of concepts are in the process of being explored, tested, and validated both in Europe and the United States. One perspective on addressing airspace structure relates to the segregation of airspace into different areas that are designed to accommodate certain types of air traffic and flows. Two such concepts are the Super Highway- a Single European Sky ATM Research (SESAR) initiative (Super Highway, 2008)- and Corridors-in-the-sky- a U.S. based approach (Hoffman & Prete, 2008). Both concepts share similarities independent of their differing operating environments in that they are designed to create special areas of airspace referred to as lanes, corridors, or tubes that aircraft with similar trajectories and in some cases similar equipment traverse through potentially high density, high complexity areas. These specialized sectors can be either static or dynamic with the overall goal of providing an area of airspace for the user that will require minimal intervention from air traffic controllers and subsequently reduce the potential workload that would exist for them without such structures.

Skippping to the third focus area of DAC, generic airspace refers to the concept of, as its name implies, airspace that has removed from its consideration many of today's necessary yet rigid components related to navigational references and sector/area specific idiosyncrasies. Through the removal of such components, it is thought that this would allow for a better utilization of workforce due to the gained ability of being able to assign controllers to virtually any sector. This is facilitated by the fact that the controller would no longer be required to know the myriad "ins and outs" of a specific sector and would be able to exercise control of the airspace through more generic methods.

This paper will focus on the second area of DAC research listed above: adaptable airspace. This area of research relates to the ability to dynamically change airspace sectorization in response to a demand-capacity imbalance without the need for altering aircraft trajectories in direct response, creating specialized areas for certain flows of traffic, or the genericizing of airspace. This dynamic sectorization would change the geometry of affected sectors to adapt to a given situation while serving to better distribute and utilize controller workload and resources.

Traditional sector geometry has been designed to take into account the relationship between traffic pattern characteristics and its effect on workload for the controller assigned that airspace. Currently, during times of demand-capacity imbalances, sectors can be split or combined depending on the needs of the situation. However, these responses often take a great deal of time and planning to implement and the decision makers typically have a limited set of options from which to choose from (Lee et al., 2008). In the Next Generation Air Transportation System (NextGen), adaptable airspace as part of DAC would allow for more strategic planning on the best sectorization responses to a given traffic situation with a more varied selection of sector geometries available based on real-time predictions and analyses. The chosen sectorizations would serve to more equitably spread and share the demand placed on a given airspace as well as the workload that would accompany that demand without the need for impacting the airspace users.

To that end, the formulation of optimal sectorizations has been an area of concerted effort through the development and testing of algorithms. A number of algorithms are in development and refinement, each with a differing approach to dynamic sectorization using different input metrics and goals to inform the calculations. A good overview and comparison of some of these algorithms can be found in Zelinsky's "A Comparison of Algorithm Generated Sectorizations" (2009). As part of the process of concept development for DAC, the development of algorithms designed for sectorization has been accompanied by a great deal of iterative fast-time modeling and testing. However, the thrust of this research has largely been limited to this area and has not had much opportunity to move forward to the next stage where the basic issues surrounding the integration of humans into the concept could be explored through exposure to algorithmically generated sectorizations in the type of environment envisioned for such operations.

The questions that guided this next phase of research into DAC came about through collaboration and consultation with fellow researchers and stakeholders. One of the central questions related to dynamic airspace reconfiguration as it relates to human-system integration is the necessary and appropriate operational procedures and guidelines to most effectively handle the transition from one configuration to another. A better understanding of the human abilities to handle such transitions and the impact that it would have on controllers is needed. Some of the fundamental questions related to airspace changes and their impact on the controllers are as follows:

1. Which airspace-related factors (e.g. airspace volume, number of aircraft affected by the boundary change, changes in the traffic flow, etc.) significantly impact the controllers during the boundary change?
2. How long does it take for the airspace transition process to complete?
3. How often can airspace be changed?
4. What factors determine when airspace change is feasible?
5. Are new automation tools required to facilitate an airspace change?
6. How much notice do controllers require prior to the boundary change?
7. What are the procedures for changing from one configuration to the next?
8. What conditions (e.g., changes in flow or complexity, increase/decrease in traffic volume, duration of peak traffic, etc.) are needed to justify a boundary change?

II. Simulation

To take a first step toward answering these questions, a human-in-the-loop (HITL) study was conducted in the Airspace Operations Laboratory (AOL) at the NASA Ames Research center from May 12-21, 2009. The intent of this study was to gain a greater understanding of the impact that dynamic airspace boundary changes have on air traffic controllers as well as to answer some basic questions regarding the feasibility and operational procedures necessary for this aspect of the DAC concept to proceed. To accomplish these goals, previous work on algorithm development was leveraged in order to provide realistic and appropriate dynamic sectorizations to the participants. Three separate algorithms were selected based upon their different approaches to sectorization and their aggressiveness related to the magnitude of change: Low, Medium, and High. The decision to choose these algorithms according to a categorization based in part on change magnitude stemmed from the exploratory nature of this study and the desire to test the range of change conditions in order to test the limits of change and more fully quantify their effects.

A. Algorithm Descriptions

For the Low magnitude change sectorizations, an approach was preferred that involved sector geometry changes that, to a certain extent, respected the composition of the current sector and would theoretically have less of an impact on the controller during a transition. This type of approach was embodied by an algorithm that

formulated sectorizations based on Dynamic Fixed Posting Areas (DFPA), which has since shifted its basis to Dynamic Airspace Units (DAU) (Klein, Rodgers, & Kaing, 2008). As with most algorithms developed for DAC, at its heart lies the end state goal of reduced workload through distribution of airspace. Because workload is an inherently subjective construct and dependent upon various factors, the algorithm must use as input proxies that can be used as estimates of what workload and complexity would amount to. Workload proxies for this algorithm are encompassed by Simplified Dynamic Density (SDD) metrics (Klein, Rogers, & Leiden, 2008), which are a subset of the larger body of metrics known as Dynamic Density (Kopardekar, Schwartz, Magyartits, & Rhodes, 2007, Kopardekar & Magyartits, 2003).

A simplified description of how this algorithm works and was initially used in generating sectorizations for this study started with a definition of the airspace that would be used as a starting point. From here the algorithm searches for shared boundary areas from which a line can be drawn that will connect the shared vertices of the two boundaries (see Figure 1). Based on this “trend” line, the algorithm begins making slices moving away from the “trend” line toward the sector’s center at 1 NM intervals. Once these slices have been apportioned they serve as the basis for a slicing series that incrementally allocates slices from one sector to another based on demand. The number of slices gained or lost by a sector depends on the calculated effect that the transfer has on reducing and distributing the SDD metrics among the sectors. Throughout the series of slices, continuity of shared areas and conformance to previous configurations is attempted such that transitions from one sectorization to the next are less extreme than they might be through a different algorithm.

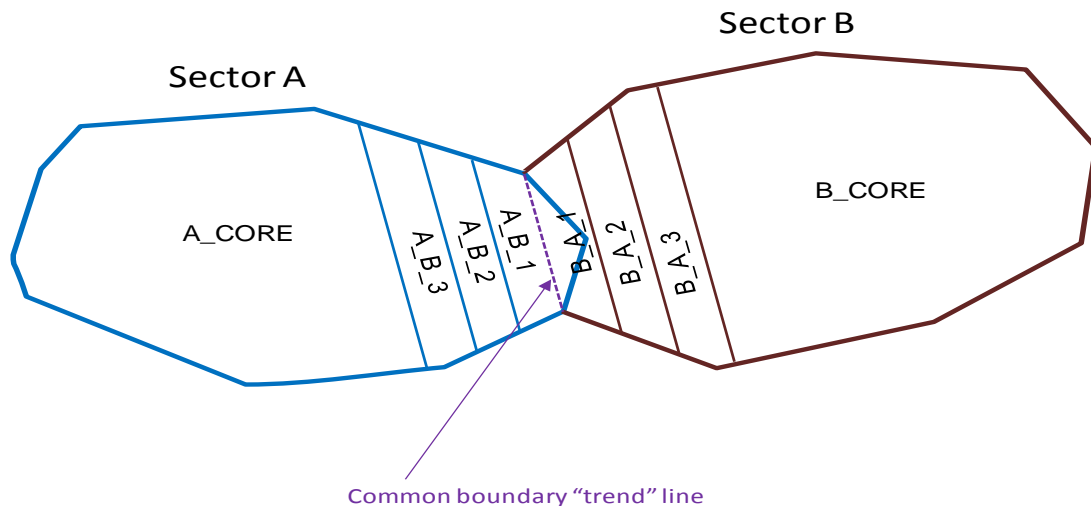


Figure 1. Example of sector slicing and resulting airspace units that can be gained or lost between sectors based on demand distribution.

For the Medium magnitude change sectorizations, a slightly more aggressive approach was desired that differed from the Low magnitude change algorithm in its input parameters and formulation. This was found in a clustering algorithm that generated sectorizations based on the grouping or clustering of flights (Brinton & Pledgie, 2008, Zelinsky, 2009). The key difference here is that in addition to incorporating DD metrics it takes into account flow information through clustering algorithms in the generation of sectorizations. This results in more drastic changes to sector geometries than the Low magnitude change algorithm because rather than focusing on boundary commonalities and how they can be configured, it makes a determination of traffic flow characteristics over time and how best to alter sector geometries to accommodate changes in those flows.

As an initial step, this algorithm calculates an independent measure of demand given the traffic for the area of concern. The flights are then clustered together according to defined clustering criteria and DD factors with an associated time dimension such that changes in clustering over time would result in changing sector geometries over time to reflect the clustering. This is done through computational geometry techniques and strives to most adequately encapsulate the identified clusters and reduce the contribution of DD factors to the complexity and workload of the controllers responsible for working the areas undergoing these transformations. Further refinements are made in an attempt to more effectively distribute DD among the sectors. An example of one proposed sectorization set based on the parameters set forth for this study can be seen in Figure 2. The left portion

of the figure shows the static, current day sector geometries for four high altitude sectors in ZKC center. Using scenario traffic data as trajectory input, the four sectors were transformed to accommodate the flows of traffic seen in the right portion of the figure in grey, overlaid on the sector boundaries.

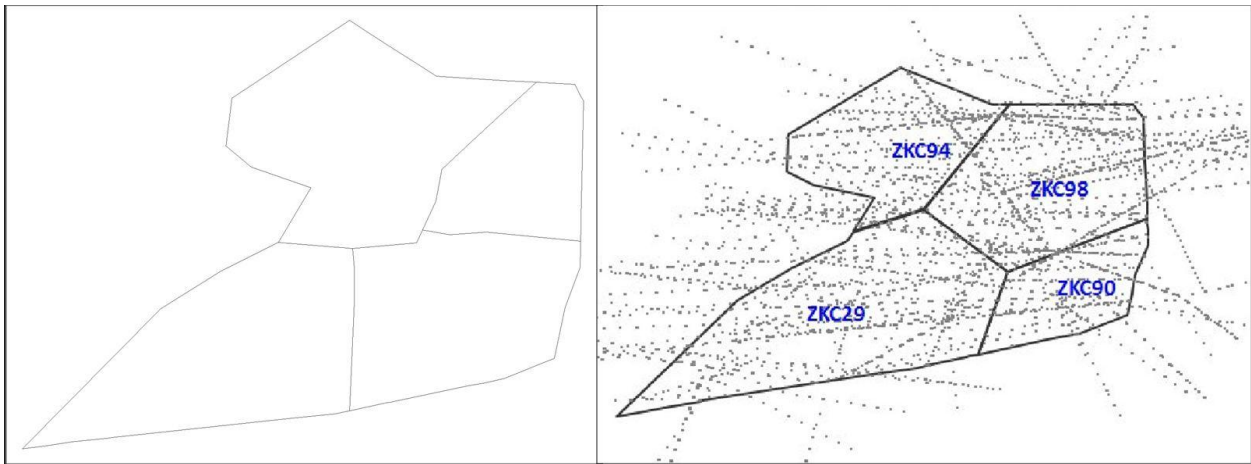


Figure 2. Example sectorization based on the clustering algorithm used for the Medium magnitude change condition.

The High magnitude change algorithm selected for this study was one based on the Mixed Integer Programming (MIP) method (Drew, 2008, Zelinsky, 2009). This particular algorithm was selected as it was observed to provide more radical sectorizations and would provide a challenging environment in which to test controller impact with. Unlike the Low and Medium change algorithms, this one did not use as input DD metrics to inform its sectorizations. Instead, it used a number of metrics (e.g. dwell time average of tracks, traffic count imbalance between sectors, boundary stability and convexity) that constituted a decidedly different approach with likewise different results.

With a predefined number of sectors to configure, the first step in reconfiguring sectors using the MIP approach involves transforming the airspace of interest into a network of hexagonal cells. Trajectory data for traffic traversing the airspace/network of cells provides the basic input for how best the cells can be connected for the most effective sector configuration. The load of each cell and the connectivity to adjacent cells- defined by the traffic flow pattern from one cell to the next- are primary considerations in the formulation of the final sectorizations. Basically, this approach does its best to respect flows of traffic without much regard to sector geometry. Due to the hexagonal partitioning of the airspace as a necessary part of the algorithm's process, resulting sectors tend to be quite jagged once the final sectors are formed. For this study, a request was made to the developers to provide additional smoothing to increase the acceptability of the sectors for the controllers without detracting too much from the intent of the final sectorizations. The compromise between the smoothing and respecting of the original sectorizations resulted in less yet still somewhat jagged sectors in the end as seen in Figure 3.

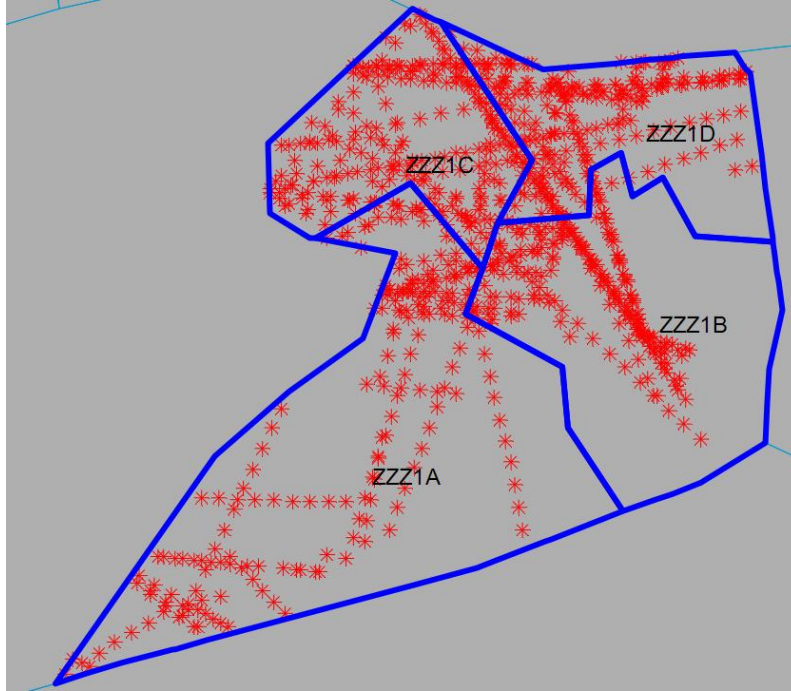


Figure 3. Example of sectorization based on the MIP approach used for the High magnitude change condition. Note the somewhat jagged edges.

B. Implementation and Approach

With the research questions defined and one method of evaluation decided through the varying of sector change magnitude -Low, Medium, and High-, a decision on the manner in which they would be integrated into a more consolidated research effort was necessary. Figure 4 shows the approach taken where a number of the questions put forth above were addressed not only through the manipulation of variables as part of the experiment design, but

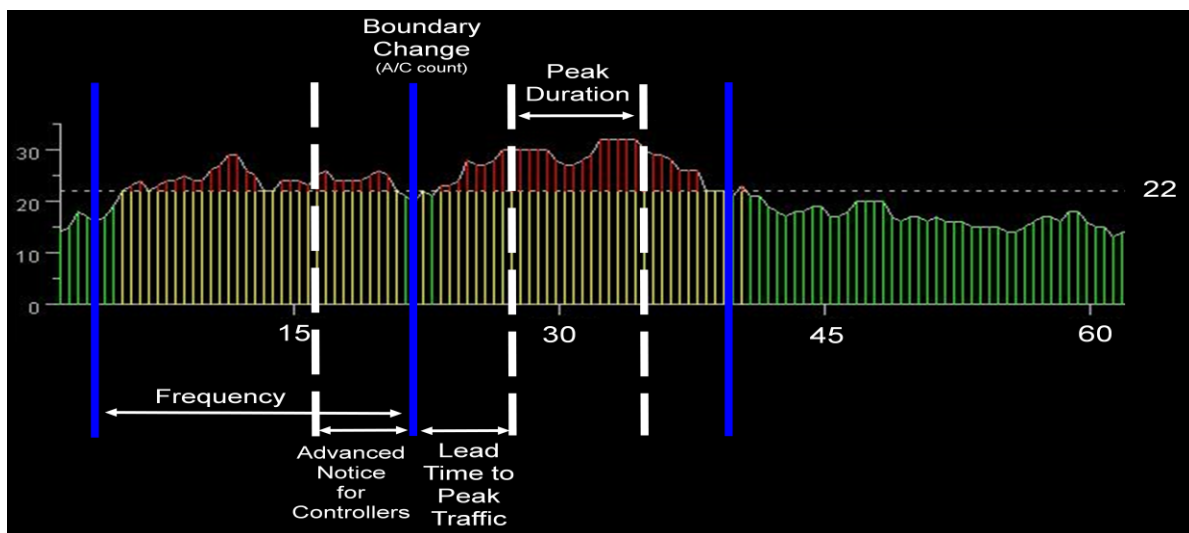


Figure 4. Traffic load in a given sector with potential boundary changes (in blue lines).

also manipulated within the design of each scenario used in testing each level of the independent variable. Without such an approach, the variables and subsequent conditions and runs necessary to address each one would quickly multiply to the point of becoming unwieldy and unrealistic to run let alone analyze.

As shown in Figure 4, some of the factors that were chosen for manipulation at the scenario level were the frequency in sector boundary change (i.e. how often), the timing of the change relative to the imposed trigger event of traffic peaks that exceed the Monitor Alert Parameter (MAP) value, and the traffic patterns that had varying durations of sustained counts in excess of the MAP value. Other factors that were intended to be manipulated in a similar fashion were standardized for the study and will be explained in the following section along with the details of the final experiment design.

III. Method

A. Experiment Design

The design for this experiment included one independent variable (IV), Magnitude of Boundary Change, which consisted of four levels: Baseline -Static, current day boundaries, Low Magnitude Change, Medium Magnitude Change, and High Magnitude Change. This was a within-subjects, repeated measures design with the four test participants being exposed to all four conditions. The dependent variables selected to measure the impact of the IV manipulations were aircraft counts on a per sector basis, subjective workload ratings, taskload, safety, and acceptability.

Factors that were varied below the level of an IV were, as mentioned earlier, the frequency of sector boundary change, the timing of the boundary change, and the traffic pattern characteristics. Factors that were held constant were the number of boundary changes, the preview time given to the controllers prior to a boundary change, and the number of sectors allowed to occupy the test airspace. The decision to hold the number of boundary changes constant was based upon the desired run length for each trial, which was 60 minutes. It was felt that any more than three boundary changes in one run would not allow for enough time in between changes to be able to assess how one change might have impacted the participants differently relative to the others without overlapping and confounding effects. With respect to the preview time given to the controllers prior to a boundary change, this factor was actually manipulated during trial runs conducted in preparation for the final study. Feedback from the participants and observation from the researchers pointed toward an optimal preview time of three minutes prior to a boundary change. This timing appeared to be just enough time to adequately prepare for the upcoming boundary change without being too far in advance as to not warrant any action. The number of sectors was limited to four and was a hard constraint imposed in the formulation of new sectorizations due to available hardware and, more importantly, the complications that would arise from swapping a number of participants in and out for each boundary change and how that type of data would be handled.

B. Participants

There were a total of four test participants. Of those, three were operations supervisors from Washington Center (ZDC), Atlanta Center (ZTL), and Indianapolis Center (ZOA), and one a recently retired controller for Oakland Center (ZOA) who actively controlled traffic within the last 4 months prior to the start of the simulation. Their air traffic control (ATC) experience spanned from 20 to 25 years with an average of 22.5 years of ATC experience.

In addition to the test participants, the duties of Area Supervisor, two Radar Associates (RAs), and “ghost” controllers responsible for all of the aircraft outside of the test airspace were performed by retired controllers from ZOA. The Area Supervisor and the two RAs played an integral role in the study. The RAs had recently retired within 2.5 and 2 years, respectively, and the Area Supervisor had retired within 6 years. All of the simulated aircraft were flown by pseudo-pilots, who were active commercial pilots and/or San Jose State University students from the aviation department.

C. Apparatus

For this simulation, the participant test area consisted of six radar controller stations (see Figure x) each equipped with 28” Barco displays, DSR keyboard and trackballs as input devices, and tablet PCs for VCS communications emulators for air-ground and ground-ground communications. In addition to these six stations were an Area Supervisor’s station with monitor and PC keyboard and mouse and a Traffic Management station configured similar to the controller stations but with extra displays for load awareness. Two side-by-side projectors were also in the test area that projected a Traffic Situation Display (TSD) with a real-time display of traffic as well



Figure 5. Test area of the AOL showing the six controller and Area Supervisor stations.

as load graphs showing the current and predicted loads for each of the test sectors. The TSD also displayed upcoming boundary previews, which was used by the Area Supervisor for controller briefings. In a separate room were the two “Ghost” confederate stations. These were equipped with 30” monitors with PC keyboards and mice. Also in another room were six pseudopilot stations which were all standard PC setups.

The common simulation platform that threads all of the positions together is the Multi-Aircraft Control System (MACS) (Prevôt, 2002). This is a JAVA based software package developed at the NASA Ames Research Center by Dr. Thomas Prevôt and his development team. MACS is able to emulate many of today’s ATC capabilities and is also scalable to test future concepts and tools. Through the development of MACS and configurations to the test stations, an environment envisioned by NextGen and specific to the testing of this DAC concept was able to be created.

D. Airspace

The test sectors used in this study were adapted from high altitude sectors in Kansas City Center (ZKC). The four sectors (ZKC sectors 94, 98, 29 and 90) are shown in green in Figure 6. Surrounding the test airspace were two “ghost” sectors, divided along East-West lines that were composed of all of the non-test airspace. Retired controllers worked the traffic in these “ghost” sectors and handled regular controller duties such as handoffs and transfer of communication (TOC) for all incoming and outgoing traffic. The flows in the test scenarios consisted of a mix of arrivals and departures to and from the area airports as well as a number of over flights en-route to various outlying destinations. For this study there were a total of six airports with arrivals and departures transitioning in the test airspace. The minimum altitude of these overflights was FL 290 with maximums being dependent upon aircraft characteristics. In general the East-West flows in these scenarios were slightly heavier than the flows running North-South.

E. Operational Environment

The operational environment for this study incorporated components of what is envisioned as part of NextGen and included some technological assumptions in keeping with that vision. With respect to assumed aircraft equipage levels, all aircraft entering the test airspace were equipped with data comm capabilities, Automatic Dependent Surveillance-Broadcast (ADS-B) reporting, and were capable of flying 4-D trajectories to a Required Navigation Performance (RNP) level of one. It was also assumed that the test airspace and all aircraft entering was conducting Trajectory Based Operations, which, for this study meant that all aircraft flying their nominal trajectories were cleared to travel according to their known route provided they remained on trajectory. This also implied that aircraft that reached their top of descent or were in a climb as a departure were cleared for their transition.

In addition to the flight-deck, a number of technological assumptions were made for ground-based capabilities as well. In this environment, the responsibility for conflict detection was shifted from the controller to ground-based

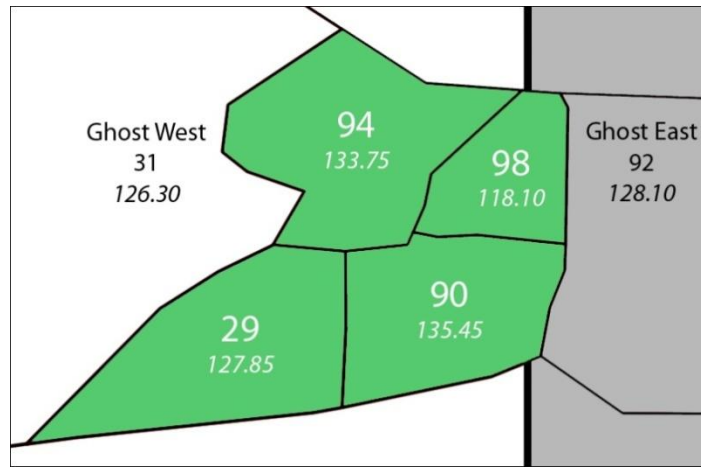


Figure 6. High altitude test airspace used in the study. The airspace was adapted from ZKC center.

automation and conflicts that were detected were conveyed to the controller through the datablocks of the involved aircraft. While the controllers always had the ability to communicate with aircraft via voice, they also had access to decision support tools that eased and expedited the process of resolving conflicts and altering trajectories. The first tool available was a trial planning function through the DSR that the controller could dynamically drag an aircraft's lateral route or trial plan a different altitude or a combination of the two that would provide real-time feedback on whether or not the proposed trajectory was conflict free. Another tool that aided conflict resolutions specifically was the auto resolver, which was the front end of an algorithm developed as part of the Advance Airspace Concept (AAC) (Erzberger, 2001). This tool provided controllers with the capability to call up, on demand, a suggested resolution for a given conflict that would be presented to the controller as a trial plan. The controller could then decide on whether or not to accept the resolution, modify it through the trial plan, or reject it and solve it without the auto resolver's support. Once a resolution was decided upon by the controller, they then had the capability to uplink the new trajectory to the aircraft through data comm. channels, which could then be loaded directly into the aircraft's FMS and flown accordingly.

F. Procedure

Aside from the formulation of the experiment's design, preparations for this study required more lead time and effort due to the involvement of and dependence on three groups responsible for developing the three sets of sectorizations based on their algorithmic approach. The first step in this process required the development of two traffic scenarios that would serve as the trajectory input for these algorithms. Scenario development was done through the scenario editor function in MACS with resulting scenarios designed for a 90 minute run length with load imbalances between the four test sectors that would drive the resectorizations in each of the algorithms. Although the actual runs were to be 60 minutes in length, longer times of sector occupancy were necessary for the algorithms to be able to generate sectorizations. The load imbalances were constructed such that they had differing characteristics between the two scenarios in terms of timing and duration and were meant to provide a basis through which comparisons could be made concerning the impact of specific boundary changes and the factors involved. The load imbalances were developed by building up traffic peaks in two of the four test sectors (ZKC 90 and ZKC 94) while keeping the loads in the other two sectors more in line with the established MAP value. A MAP value of 22 was selected for each of the sectors in the test airspace and the peaks were built to reach counts of up to approximately 28 aircraft in each of the two targeted sectors. In each of the two scenarios, two peaks were built in at different times and the duration of time that aircraft counts exceeded the MAP value also differed.

After completion of the two test scenarios, they were run in the AOL where trajectory data could be collected and sent to the three algorithm development teams for testing and sector generation. The constraints placed on the generation of the sectorizations were that the initial sectors always needed to start with the current day configuration, there would always be four sectors, the outer boundaries of the overall test airspace encompassing the four sectors would remain constant, and that the altitude dimension would also remain constant. The Low magnitude change algorithm resulted in revised sector configurations at 15 minute intervals for the two traffic scenarios that covered the 90 minutes of trajectory data. This resulted in six sectorizations for each scenario. The Medium and High magnitude change algorithms generated sectorizations at five minute intervals based on a 30 minute look ahead of trajectory data. This resulted in 18 sectorizations per scenario from which to choose.

After the traffic scenarios and sectorizations were complete, final decisions were made on when the three boundary changes would occur in each run relative to traffic peaks. Prior to this point, it was decided that there would be a total of four change conditions, each with different boundary change points. This meant that replications of each of the two scenarios were made, totaling four final scenarios, with slight variations to each so as to appear different enough to the participants yet remain true to the original intent of the traffic patterns. The results of the boundary change decisions and the final traffic patterns can be seen in Figure 7.

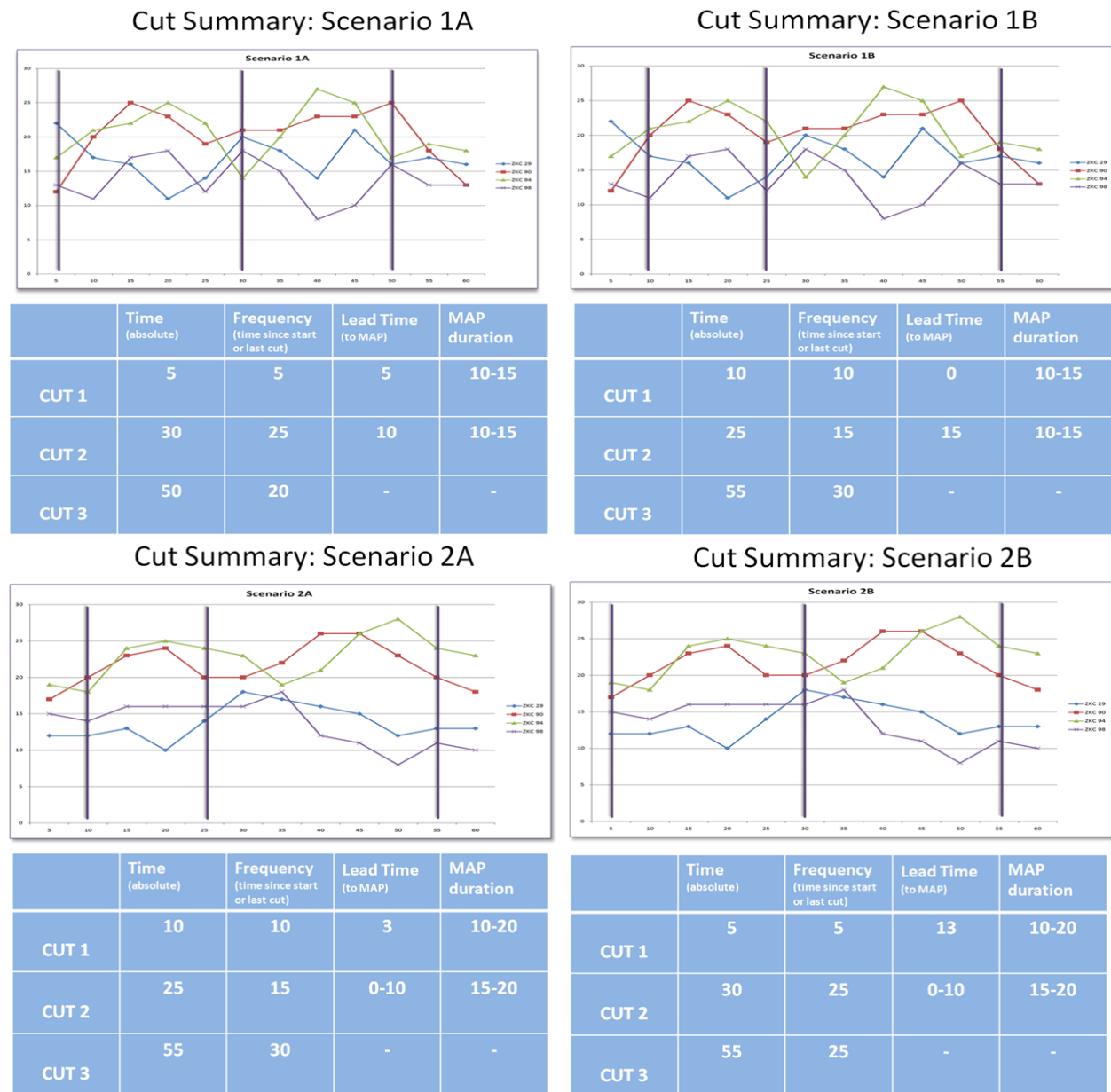


Figure 7. Summary of final boundary change points and traffic load characteristics.

Having decided upon the boundary transition points, the next step was to match the different sectorizations to the times. This meant that three sectorizations were mapped to each of the four boundary change conditions resulting in a total of 12 sector configurations per algorithm that would be used for the study. While the finalization of boundary configurations was taking place, trial runs were being conducted in the AOL using local retired controllers from ZOA center. Through these runs a number of procedures, boundary change, and display issues were examined and tested. One of the results of these trial runs was the preview time needed for the controllers prior to a boundary

change. Feedback from the controllers was that three minutes was nominally enough time to handle the tasks associated with a boundary change. Another important piece of input from these trial runs was in response to the sectorizations provided for the Low magnitude change condition. This condition was intended to reflect a more human-centered design to sector configuration and feedback was desired from the subject matter experts regarding the current design of the provided sectors. This was obtained through the presentation of each of the boundary configurations received for the Low magnitude change condition in a post-session questionnaire packet along with questions regarding some of the factors that impact the acceptability of the initial sector configurations. Following the completion of all trial runs and adjustments made to the necessary boundary configurations, the specific sectorizations were organized and prepared for final implementation. With the exception of the Low magnitude change, the final sectorizations were left untouched, true to their original intent. Further modifications were possible, but in order to test the full range of impact of the designs, sector characteristics such as jagged edges and extreme changes in the geometry and relative positions of sectors were allowed to remain. This also served as a benefit to algorithm developers in being able to identify additional considerations for their respective approaches. Figure 8 presents an example of the first boundary change from the Low, Medium, and High magnitude change conditions for one of the four final boundary scenario conditions (1A in Figure 7). Note that the initial sector configuration, regardless of condition, was the current day sector configuration as presented in Figure 6.

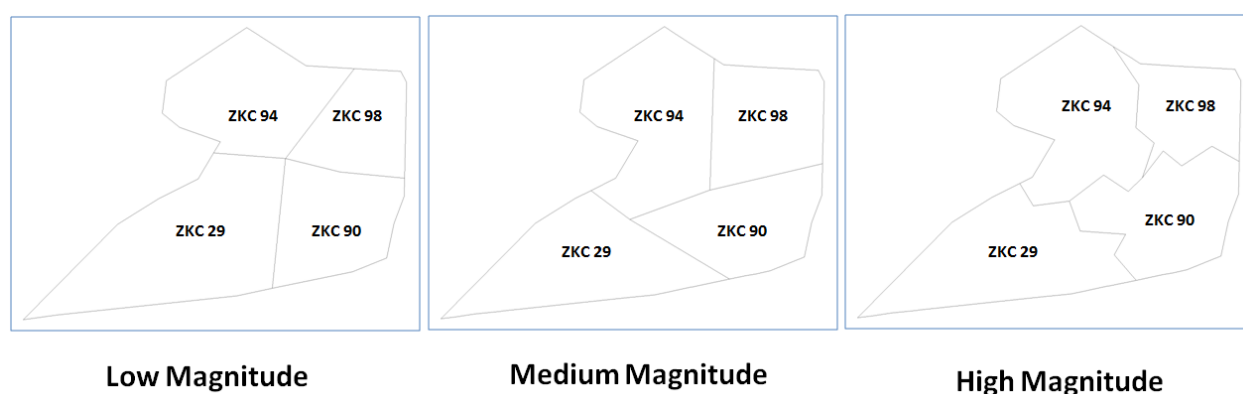


Figure 8. Example of the first sector boundary change for each of the change conditions in Scenario 1A.

This initial HITL study was conducted over the course of two weeks and involved the participation of four active operations supervisors as R-side controllers for the four test sectors, two recently retired confederate controllers as supporting Radar Associate (RA) controllers, and a recently retired confederate controller acting as an area supervisor. Two additional retired confederate controllers acted as “ghost” controllers that handled air traffic feeding into and exiting the test airspace. After the initial briefing and administrative duties were performed, training on the tools and environment was started in the AOL. The training was conducted over the course of three days and involved the use of training traffic scenarios that started with 50% of the eventual traffic levels that would be presented during the data collection runs, and progressed to 70% and eventually 100% of the traffic levels once proficiency with the tools and understanding of the airspace was observed. The entire first day of training was devoted to tools and airspace familiarization with the first set of boundary changes not being presented until the second day. The sectorizations used for the training were selected from each of the three different algorithms with consideration to not include those that would actually be used during data collection.

Data collection began on the fourth day after an initial briefing and discussion. The run order was counterbalanced in order to minimize the possibility of confounding data related to order and training effects. A total of 16 runs were conducted with each of the four boundary change conditions (Baseline, Low, Medium, and High magnitude change) being presented four times. Throughout each of the runs, real-time workload ratings were collected through an integrated workload assessment keypad in MACS on a scale from one to seven, with one being the lowest and seven the highest rating of perceived workload. Prompts were presented at different times according to the boundary change schedule in each of the four scenario types used. The basic sequence of workload prompts was three minutes and one minute prior to a boundary change, and likewise following the change. With boundary changes more than 10 minutes apart, workload prompts were presented at five minute intervals until the next change. For Baseline runs, workload prompt sequences were identical to the ones used in each of the corresponding

four scenario types for the runs involving boundary changes. This was done in an effort to isolate workload differences that were attributable to the boundary change rather than to specific characteristics of the traffic pattern.

For each run, the four test participants were, similar to today, responsible for making and taking handoffs of aircraft entering and exiting their sector and maintaining safe separation distances in accordance with current separation minima. The Area Supervisor monitored the current and predicted traffic loads for the test sectors through interactive load tables and graphs, and assessed the workload of the controllers. For non-Baseline runs, the Area Supervisor also had access to a preview of each upcoming boundary change. Based on the supervisor's judgment of available resources and the work that would be required before and after a particular boundary change, RAs were assigned to assist one of the R-side controllers. Due to the layout of the lab in this study, there were two RA positions situated between an R-side pair on opposite sides of the room. Because of this layout, support from the RA controllers was limited to only one of the associated R-side pairs. The pairs were divided according to horizontal proximity of the initial sectors such that ZKC 29 and ZKC 90 was a pair assigned one RA controller and ZKC 94 and ZKC 98 were assigned the other. The RA position assisted one position at a time and had access to the same tools and functionality that the R-side position had. Prior to being assigned a position to support, the Area Supervisor would gather the individuals for a quick brief on the current or upcoming situation. Following this brief and after the supervisor gave the final go-ahead, a lab team member would activate the appropriate RA position.

For boundary change conditions, the controllers on station would receive a boundary preview three minutes prior to the change. Figure 9 presents what this preview looked like to the controllers. The displays had a small boundary preview window that showed a countdown until the boundary change (this window was always displayed) and the upcoming boundary would be overlaid on the current boundary. Additionally, the corresponding sector numbers were also displayed with the preview for the controllers to have an awareness of the new configuration. The sector numbers would also remain on the display for one minute following the boundary change to assist with handoffs and general situation awareness. Typically upon activation of the change preview, the R-sides and RA controllers, if applicable, would begin making handoffs and point outs to the appropriate sectors. Transfers of communication were automated so these tasks were done through the DSR keyboard. Some point outs were done verbally. If all steps proceeded ideally, all aircraft would have been transferred to the appropriate sectors and under their track control at the time of the actual boundary change. If some boundary changes appeared particularly workload intensive the supervisor would often assign an RA to assist with the initiation of handoffs and switch position assignments shortly after to assist the receiving sector in managing the new set of aircraft gained.

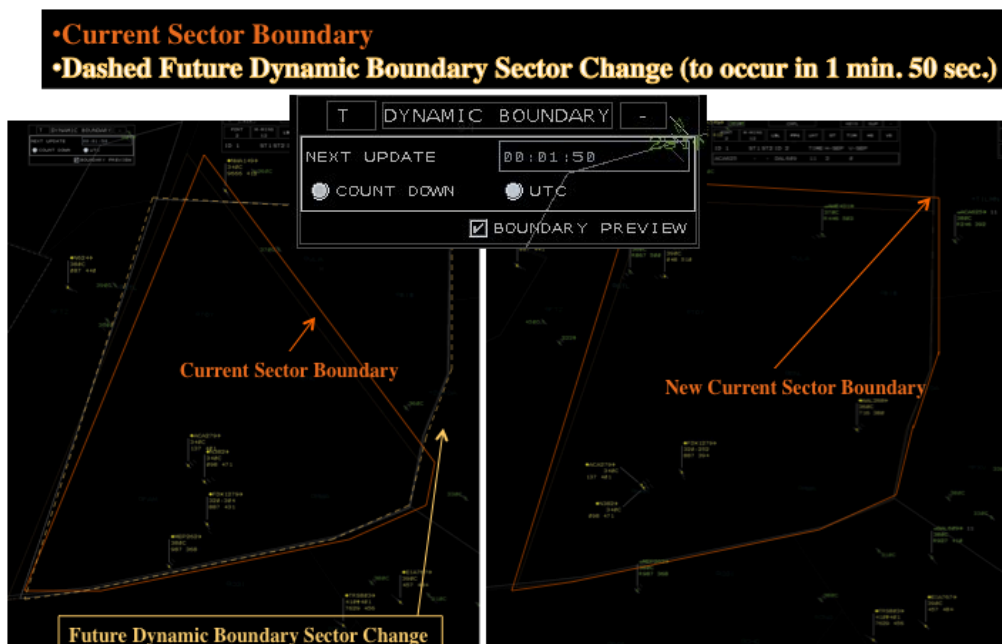


Figure 9. Example of a sector boundary preview (left) and the actual change (right). A countdown was also provided indicating time left until next boundary change.

Following each run, a post-run questionnaire was distributed to each participant. This questionnaire asked questions regarding overall workload experienced during the previous run, acceptability of boundary changes if applicable,

and solicited comments on what factors affected their acceptability ratings for the boundary changes in that run. In addition to this data, a wide variety of other data related to workload, aircraft trajectories, system interaction, safety, among other things were recorded through the data collection system of MACS. Screen recordings of all stations were also made for each run for reference later during analyses. Another source of data came from observers that sat next to the controllers during the run taking note of communications and coordination between sectors and any special situations that arose throughout the course of the run. A post-simulation questionnaire was given following the final data collection run followed by a debrief discussion involving all participants where feedback regarding the concept, tools, automation, and a host of other topics were covered and discussed collaboratively.

IV. Results

Analysis of the volume of data collected during this study is ongoing. The results presented in this section relate more to an overall comparison of the four boundary change conditions- Baseline, Low, Medium, and High magnitude change- than to an investigation into the specific factors tested by the individual boundary changes in each run. This latter analysis will be forthcoming. Results to be presented here will pertain to quantifying the change in airspace for each of the applicable boundary change conditions, traffic characteristics, taskload, workload, safety as it relates to conflicts and operational errors, and finally the acceptability of the different approaches to airspace configuration.

A. Airspace

The first aspect of airspace to be evaluated is the percentage of airspace volume that was gained and lost through the boundary changes in each of the boundary change conditions. The percentage values with regard to volume gained refer to how much new airspace, in addition to the previous sector, was gained upon the reconfiguration. The percentage lost refers to the volume of airspace previously assigned the sector prior to the change but is no longer assigned after the change. The two separate measures, therefore, allow for a sector to both gain and lose airspace volume through a boundary change. The analysis of this data serves as a validation of the categorizations applied to the different sectorizations provided by each of the three algorithms and from which subsequent analyses will refer to. Figure 10 presents the descriptive data for both of these metrics.

In terms of airspace gained, the results support the categorizations of magnitude change that were applied. The mean percentages of airspace volume gained for the Low magnitude change conditions was 9.56 ($SD = 3.58$), the Medium change condition had a higher mean percentage of 15.99 ($SD = 10.35$), and the High change condition resulted in the greatest percentage of volume gained with a mean value of 23.66 ($SD = 7.12$). A comparison of the means was conducted with a one way Analysis of Variance (ANOVA) where a significant main effect was found with a confidence interval of .05 ($F(2,33) = 10.52$, $p < .01$). A Bonferroni post hoc test was conducted to further investigate the differences between the means, which showed a significant difference between the percentage of volume gained in the Low and High magnitude change conditions ($p < .01$) and a strong yet non-significant difference between Medium and High ($p = .054$).

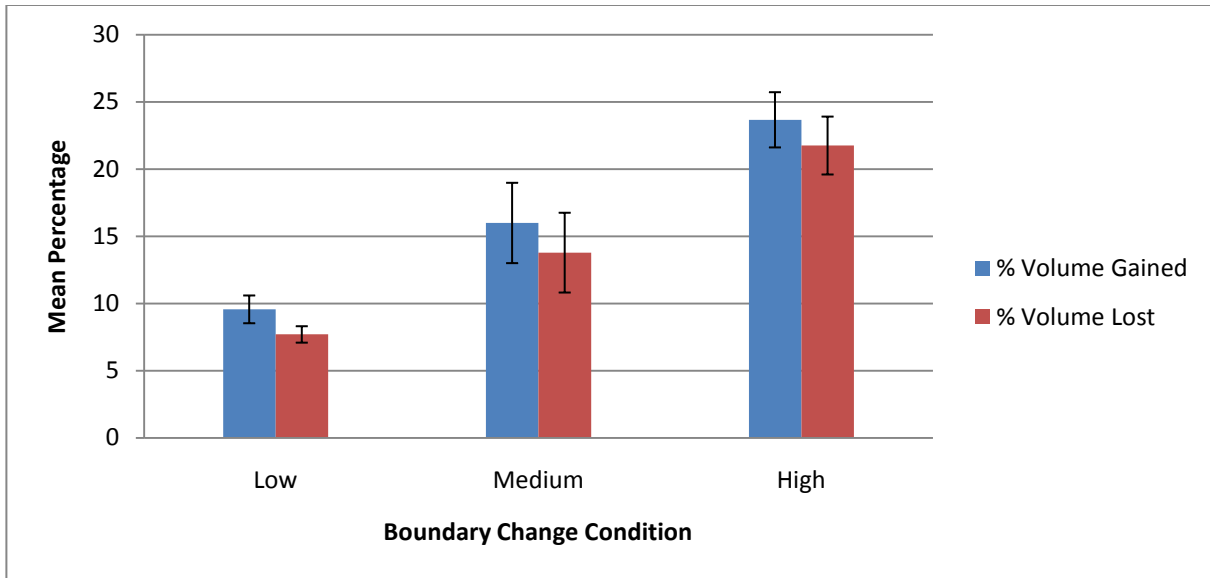


Figure 10. Differences in airspace volume gained and lost between boundary change conditions.

Results for airspace volume lost were similar to those observed for gain. The Low change condition had a mean percentage of volume lost of 7.69 ($SD = 2.11$), the Medium change condition had a higher mean percentage lost of 13.78 ($SD = 10.28$), and the High change condition had a mean loss of 21.75 ($SD = 7.46$). To investigate the differences between these means a one way ANOVA was conducted where a main effect was found ($F(2,33) = 10.79$, $p < .01$). A Bonferroni post hoc test was conducted to investigate the significant main effect further and significant differences were found between the Low and High change percentages of volume lost ($p < .01$) as well as the Medium and High change conditions ($p < .05$).

B. Traffic Characteristics

1. Traffic Distribution

The results presented in this sub-section refer to the impact that the different boundary change conditions had on the distribution of traffic among the sectors as well as the effect that the distribution had on managing loads to the established MAP value and the dwell times for aircraft in the test sectors. The first result of traffic distribution among sectors is presented in descriptive terms and represents the difference of the standard deviations of aircraft count between the four test sectors over time for each of the boundary change conditions. Because the initial analyses of data was focused more on making overall comparisons, the aircraft counts for the test sectors were collapsed across the four scenario types for each of the four boundary change conditions (Baseline, Low, Medium, and High). This only affords a very general comparison but provides an opportunity to see how each of the boundary change conditions managed the distribution of traffic across four different situations. The collapsing of data was done by averaging the aircraft count for each of the four test sectors across the four scenario types. This was done for each boundary change condition. Following this step, the standard deviations of the aircraft counts for the four test sectors were then calculated for each of the 60 minutes. In relative terms, results showing lower values for standard deviation are interpreted as a particular boundary condition being more effective at distributing aircraft among the sectors because the differences in count between them are less. Figure 11 presents the results where it can be seen that the earlier time segment with the initial traffic peak has a tighter clustering between each of the conditions with the Low magnitude change condition having the lowest amount of deviation followed by the High magnitude change condition. However, in the latter half where there was a more sustained traffic peak, the boundary change conditions begin to differentiate a bit more. In particular, the Baseline and Low magnitude change condition show higher deviation among the sectors whereas the Medium and High magnitude change conditions showed lower deviation suggesting that the greater magnitude in change in these sectorizations resulted in a more even distribution of aircraft among the test sectors. Interestingly it appears as though, overall, the High magnitude change condition managed to distribute the traffic count more equitably than the other change conditions.

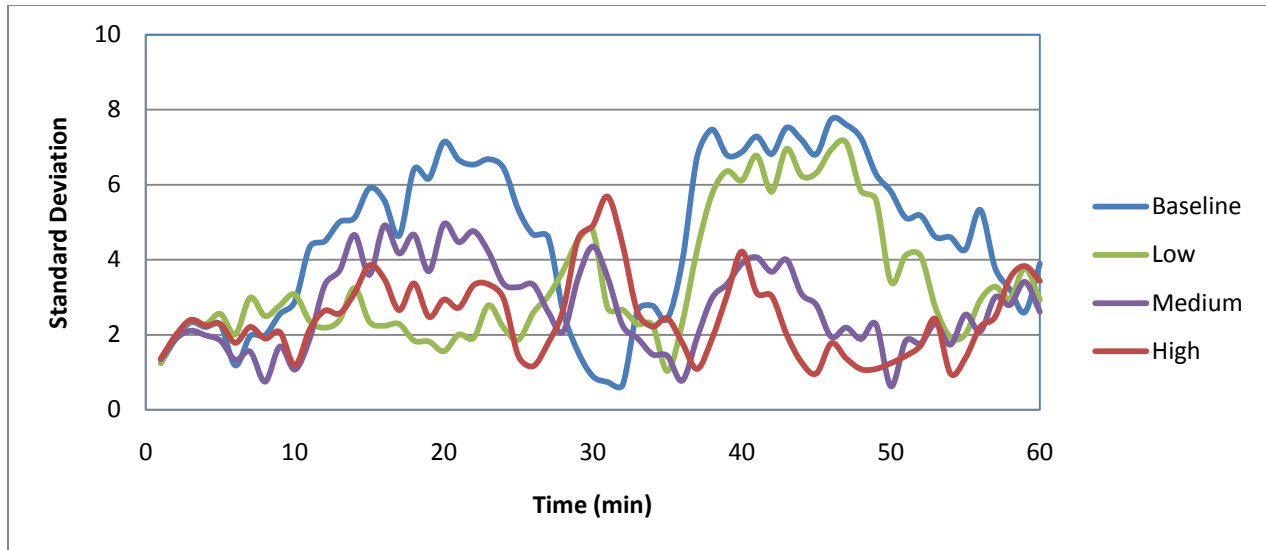


Figure 11. Standard deviation of aircraft count among the four test sectors plotted over time for each boundary change condition.

2. MAP Value Management

The results presented here refer to the amount of time in each sector that the aircraft count exceeded the MAP value of 22 aircraft. This is another indication of how well each of the different approaches to sectorization performed the task of traffic distribution. For this analysis, the total number of minutes above MAP in all four test sectors was initially summed and averaged for each run with the final mean across the four scenario types being reported and compared here. Before this analysis was conducted, it was assumed that the Baseline condition, without any boundary changes, would show the greatest mean duration over MAP compared to any of the other conditions. And this was supported (see Figure 12) in that the mean duration for Baseline was 14.63 minutes ($SD= 0.75$), whereas the Low magnitude change condition had a mean duration of 11.25 minutes ($SD= 3.40$) over MAP, the Medium magnitude change condition a mean of 10.50 minutes ($SD= 3.01$) over MAP, and the High magnitude change condition resulting in the least amount of time over MAP with a mean of 8.50 ($SD= 2.03$). A comparison of the means was conducted through a repeated measures one way ANOVA where a significant main effect was found, $F(3, 12)= 4.12$, $p< .05$. This meant that there was a significant difference between the boundary change conditions. Further analysis into the differences between the means showed that the only comparison to have a significant difference was between the Baseline and High magnitude change conditions ($p< .05$). This meant that the High magnitude change condition was significantly better than Baseline in keeping counts below MAP whereas the other conditions were not. It is likely that the reason for the non-significant findings between the other comparisons was due to the standard deviations of the non-Baseline runs being rather high.

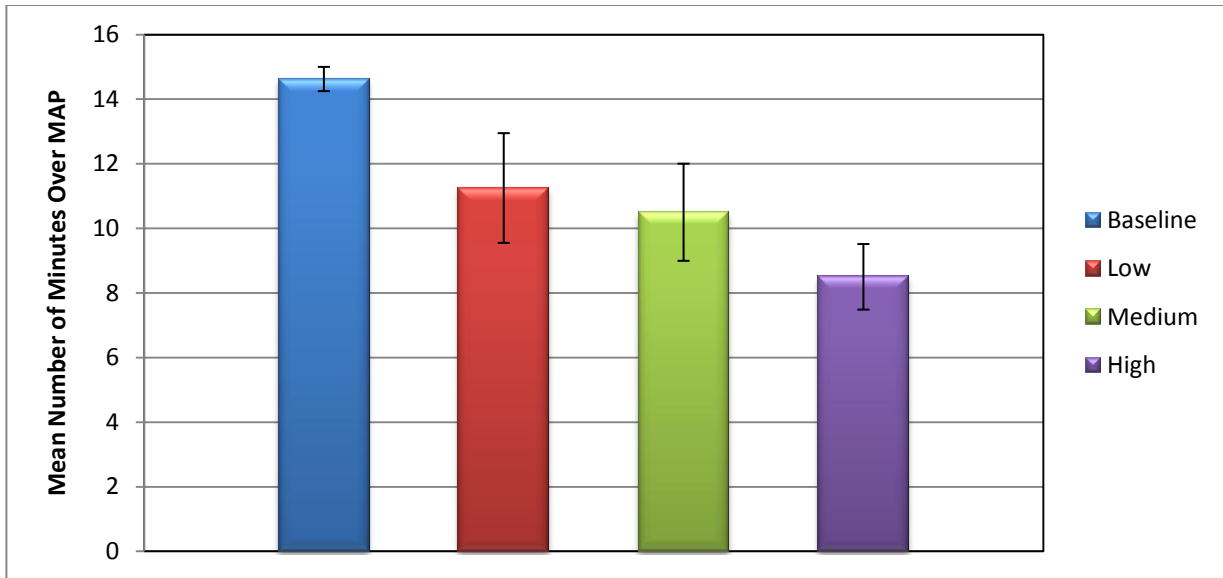


Figure 12. Mean duration of minutes in excess of the MAP value across all four test sectors.

3. Aircraft Dwell Time

Another measure of the impact that sectorizations might have on traffic and associated workload is the amount of time spent in a sector. The longer an aircraft spends in a sector, in geographic terms, the more stable it is and the more time the controller has to incorporate that aircraft into his or her operational picture. To address this factor, the average time spent in a sector was examined for all four boundary change conditions. The descriptive statistic for this measure as well as Figure 13 show that the Baseline condition had the greatest mean dwell time ($M= 462.12$ seconds, $SD= 75.87$) followed in decreasing order by the Low magnitude change condition ($M= 439.14$ seconds, $SD= 61.18$), the Medium magnitude change condition ($M= 424.69$ seconds, $SD= 41.25$), and finally the High magnitude change condition with the shortest mean transit time of all conditions ($M= 399.22$ seconds, $SD= 53.29$). A repeated measures one way ANOVA was used to examine the differences in these means and a significant difference was found, $F(3, 60)= 3.167$, $p< .05$. Further inquiry using a Bonferroni post hoc test revealed a significant difference in mean dwell time between the Baseline and High magnitude change condition ($p< .05$). This meant that the sectorizations encountered in the High magnitude change condition resulted in aircraft travelling through sectors for a significantly shorter duration of time, which is an indication of less stability in the sectorizations and potentially greater difficulty in handling aircraft and maintaining a coherent picture of the traffic in this condition.

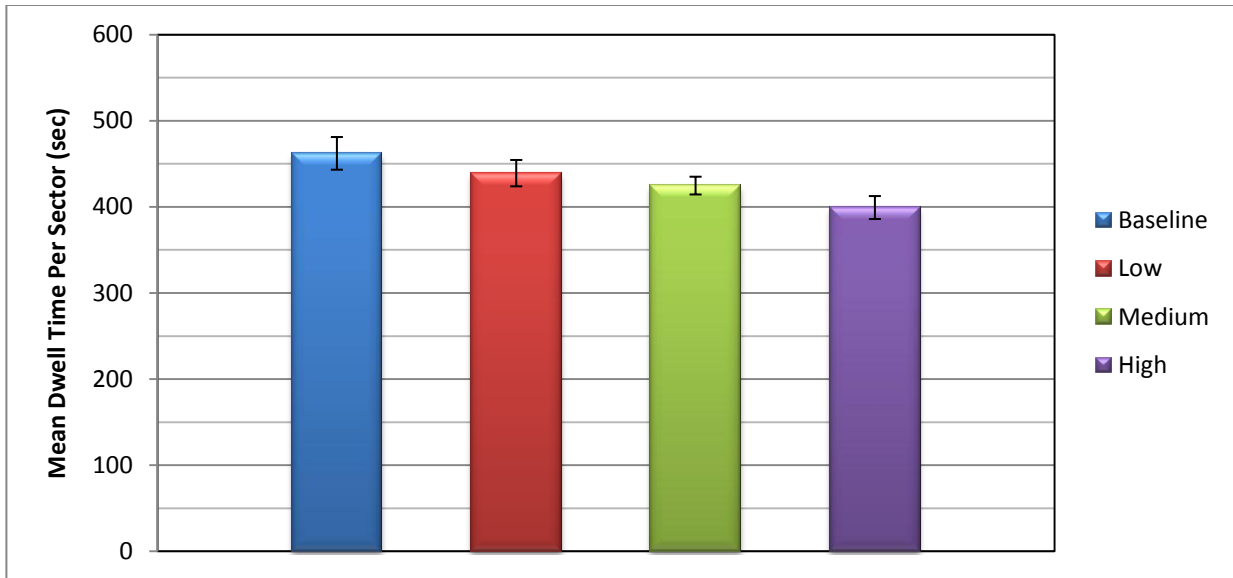


Figure 13. Mean dwell time of aircraft in the four test sectors. A lower value is considered worse in terms of stability and situation awareness.

C. Taskload

1. Handoffs

In the analysis of taskload, comparisons will be made between the four boundary change conditions that deal with controller initiated actions as they relate to the number of handoffs and point outs performed. These results will provide insight into the difficulty imposed by the various sectorization approaches. The first measure of task load to be analyzed is the number of handoffs that were required in each of the conditions. The analysis was done by averaging the total number of handoffs performed across the four scenario types for each of the boundary change conditions. Figure 14 shows that the Baseline condition had the fewest ($M= 475.75$ handoffs, $SD= 8.69$) with the next highest number of handoffs being the Low magnitude change condition ($M= 500.25$ handoffs, $SD= 22.47$). This trend continued with the Medium change condition having the next highest mean of 513.75 handoffs ($SD= 21.75$) and the High change condition requiring the highest number of handoffs to be performed ($M= 522.25$ handoffs, $SD= 26.61$). A one way repeated measures ANOVA was run on these means where a significant difference was found, $F(3, 12)= 3.74$, $p< .05$. However, a Bonferroni post hoc test did not reveal a significant difference between the boundary change conditions, although the significance value between the Baseline and High magnitude change conditions was equal to .052.

2. Point Outs

Another measure of the taskload required for boundary changes is the number of point outs. This refers to cases where certain aircraft may not enter a proximal sector's airspace but come close and thus require notification and awareness. This task requires extra time and coordination and adds to the taskload and workload of a controller. It should be noted the numbers presented here are a minimal subset of all point outs performed due to the fact that while the participants made a great effort to perform point outs through the DSR keyboard, a number were performed verbally and those counts were missed. This data should be used in support of the handoff results in highlighting the progressive difficulty imposed by the higher degree of airspace transition. Figure 15 presents the descriptive statistics on the available data, which shows that, similar to handoffs, the Baseline condition required the fewest number of point outs ($M= 38.25$, $SD= 3.78$), the Low magnitude change condition required the next highest number ($M= 48.50$ point outs, $SD= 13.48$), the Medium change condition had the next highest number of point outs ($M= 56.25$ handoffs, $SD= 21.42$), and the High change condition showing the highest number of point outs ($M= 68.25$ handoffs, $SD= 15.31$). A one way, repeated measures ANOVA was conducted to further examine the differences between these means, but the results were not significant, $F(3, 12)= 2.88$, $p> .05$. Despite this result, the increasing trend is clear and in line with the handoff results.

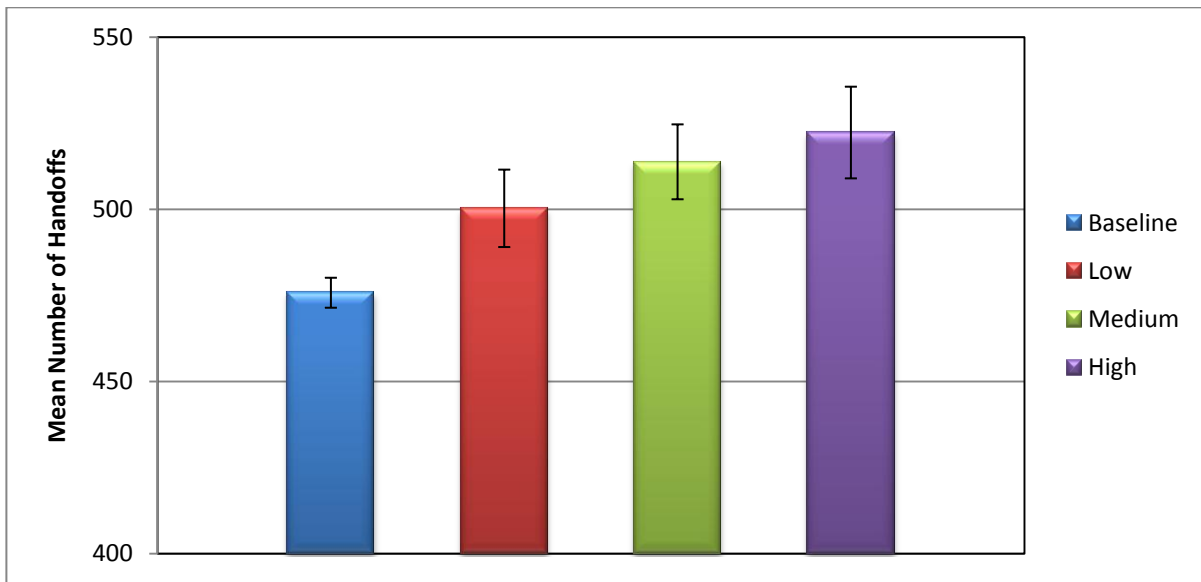


Figure 14. Comparison of mean number of handoffs performed for each of the boundary conditions.

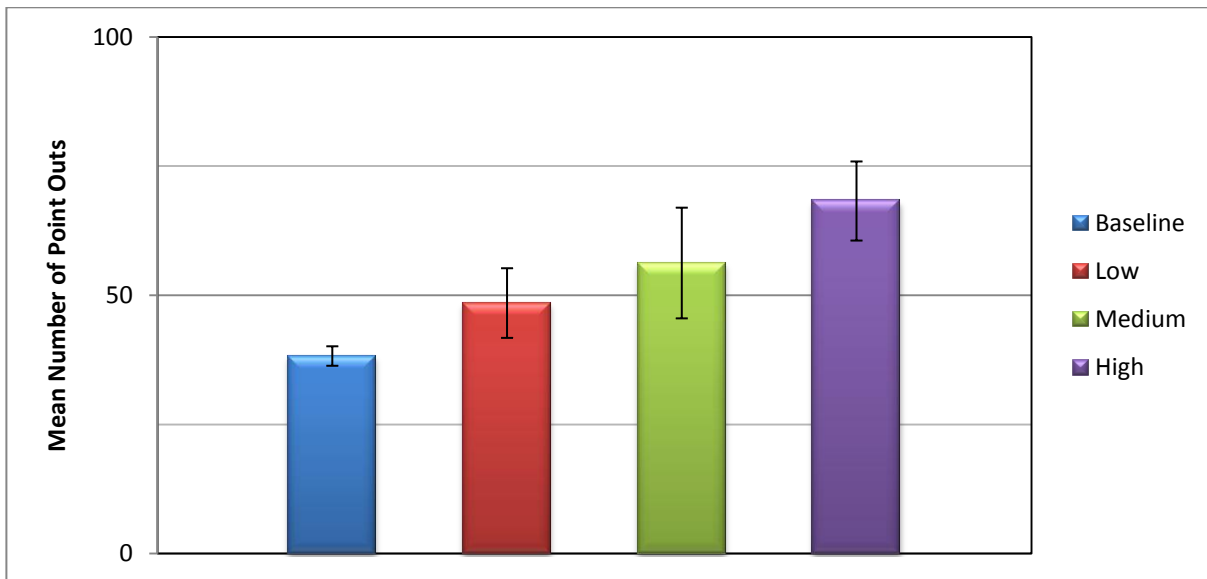


Figure 15. Mean number of recorded point outs performed for each boundary change condition.

D. Workload

Workload was obtained through prompts displayed on the DSR display in MACS throughout each run. Participants rated their perceived workload based on a 7-scale rating with one being the lowest possible rating and 7 the highest. The sequences of prompts were at different times in each run according to the boundary change point. Prompts were at three minutes and one minute prior to and likewise after the change with prompts every five minutes during periods of 10 minutes or more between boundary changes. The workload results presented here will relate to the average workload reported for each run as well as the average workload limited to the time segments surrounding the boundary changes only.

1. Mean Run Workload

The mean workload for a run takes a more general view of the workload that was experienced in each of the boundary change conditions taking into account the time in between boundary changes. Figure 16 presents the descriptive results where it can be seen that there is a consistent upward trend in reported workload along with the boundary change condition. The mean workload reported in the Baseline condition across the four scenario types was the lowest at 4.40 ($SD= 1.07$). Results for the Low change condition were higher at 4.85 ($SD= 0.98$) followed by the Medium change condition at 5.02 ($SD= 0.88$) and the High change condition with the highest mean workload rating of 5.45 ($SD= 0.49$). Looking at these descriptive statistics it can be seen that in addition to the increasing workload reported for the boundary conditions, the standard deviations decrease suggesting that workload becomes increasingly greater for all of the sectors instead of the disparities evident in the Baseline condition. To look more closely at the differences in mean workload, a one way, repeated measures ANOVA was conducted where a significant difference was found $F(3, 12)= 3.74, p< .05$. Similar to previous results, a Bonferroni post hoc test revealed that there was a significant difference between the Baseline and High magnitude change condition in mean reported workload ($p< .01$) but no significance for other comparisons.

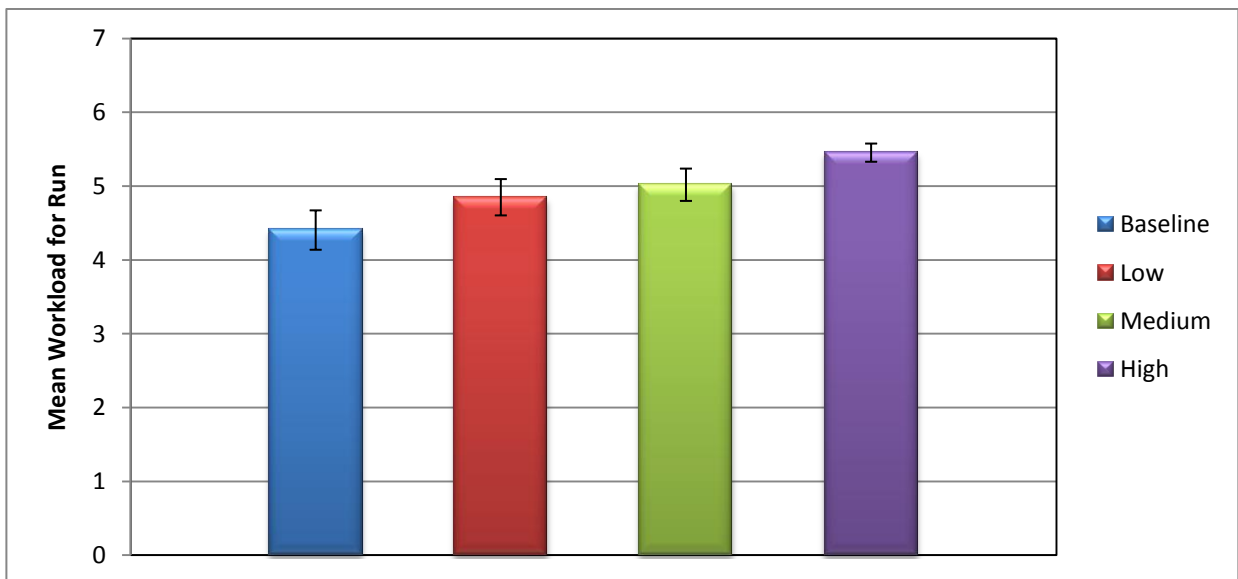


Figure 16. Mean workload reported throughout the run. Note the reverse trend in standard deviation.

2. Workload at the Boundary Change

The workload used in this portion of the analysis was limited to the ratings observed at the three and one minute prompts on either side of the boundary change. It was thought this would remove the extraneous factors unrelated to boundary changes and therefore provide a clearer picture of the workload as a function of the boundary change condition. The descriptive statistics for this metric show that, similar to overall workload, the Low magnitude change condition resulted in lower workload relative to the other change conditions with a mean of 4.75 ($SD= 1.48$). The Medium change condition resulted in a mean workload rating of 5.06 ($SD= 1.45$). The High change condition again resulted in the highest workload rating with a mean of 5.63 ($SD= 1.14$). Relative to the overall workload reported above, limiting the results analyzed to the boundary change shows that the mean workload for the Low change condition actually decreased by doing this while it increased for the Medium and High change conditions. The High change condition resulted in the greatest increase by this limitation meaning that the boundary changes in this condition had a greater impact on workload than an overall view would initially suggest. Results from a one way, repeated measures ANOVA show a significant difference between the means, $F(2, 141)= 5.05, p< .01$. Bonferroni post hoc tests also showed a significant difference between the Low and High magnitude change conditions ($p< .01$).

E. Safety

The results for safety will cover three areas: number of conflicts, airspace deviations, and losses of separation. Each of these metrics has its own implications on safety with conflicts relating to airspace complexity and potential

losses of separation, and airspace deviations and losses of separation falling within the category of operational errors deserving of investigation per incident.

1. Number of Conflicts

The number of conflicts refers to the number of unique instances of each pair of aircraft detected by the conflict probe within the test area to lose separation. Excluding false alerts, each of these instances required an action by the controller so it is at once an indicator of safety as well as one for taskload and a contributor to workload as well. Descriptive statistics (see Figure 17) for this metric show that there was approximately the same mean number of conflicts per boundary condition with Baseline having a mean number of 70.00 conflicts ($SD= 6.63$), the Low change condition had a mean number of 64.25 conflicts ($SD= 13.60$), the Medium change condition had a mean number of 62.00 conflicts ($SD= 7.87$), and the High change condition having a mean number of 66.75 conflicts ($SD= 12.23$). Conducting a one way, repeated measures ANOVA on these means did not reveal a significant difference, $F(3, 12)= 0.43, p> .05$.

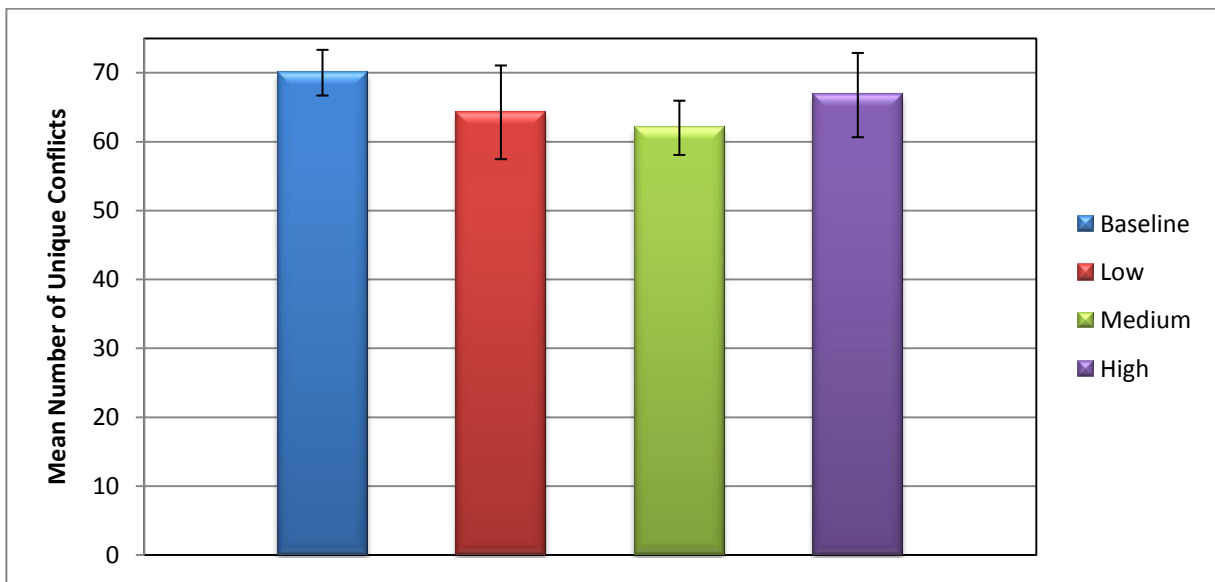


Figure 17. Mean number of unique conflict occurrences per boundary change condition.

2. Airspace Deviations

This metric was broken down into three different components, each representing a different case where an aircraft was in violation of a particular sector's airspace. Results relate to the difficulty in dealing with traffic situations represented by Baseline, and the impact of certain boundary change types on the ability to deal with the situation as well. The components that this topic was broken down into are handoffs initiated late (Late init HO), which is a more normal case where an aircraft is owned by one sector and has entered another sector's airspace without a handoff. Another case is a late handoff bypass (Late init HO Bypass) where an aircraft is in another's sector without a handoff, and then is handed off to a sector further downstream that bypasses the sector of ownership. Another category of deviations involves the acceptance of handoffs, referring to the late acceptance of a handoff (Late Accept HO) (i.e. an aircraft is in a sector with an initiated handoff but is not accepted until later). An important note here is that for this study the normal rules of "spinning" an aircraft if a handoff is not accepted were not in effect. Figure 18 presents the total counts across the four boundary change conditions broken down by type of airspace deviation.

The next step in this analysis was to take the total number of deviations for each of the scenario types and average those values for each boundary change condition in order to make a comparison of the means. For the Baseline condition, a mean number of 103.50 deviations ($SD= 17.02$) occurred, with an increasing trend observed starting with the Low change condition ($M= 152.00, SD= 56.49$) continuing with the Medium change condition ($M= 155.50, SD= 27.19$) and the High change condition showing the largest increase in airspace deviations relative to the other change conditions ($M= 220.00, SD= 16.59$). A comparison of the means shows a significant difference through a one way, repeated measures ANOVA ($F(3, 12)= 8.13, p< .01$). Similar to other analyses, a significant difference between the Baseline and High change condition was observed ($p< .01$).

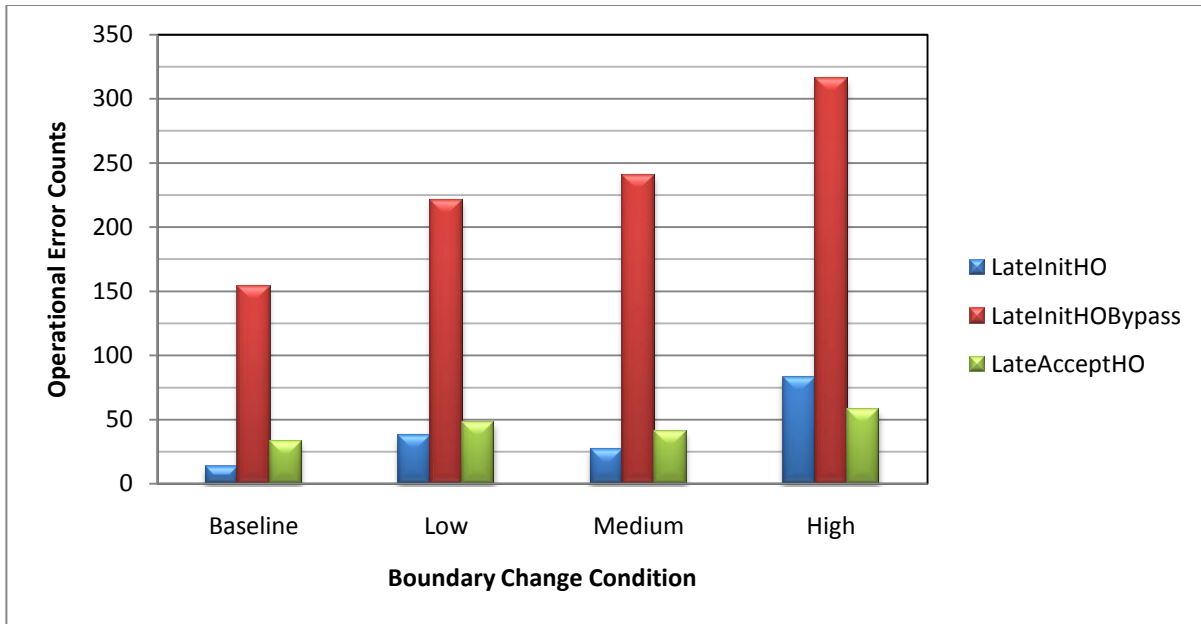


Figure 18. Total counts of airspace deviations by deviation type.

3. Losses of Separation

This category is the most serious as it represents cases where two aircraft violated the separation minima defined by five NM of lateral separation and 1000 feet of vertical separation. An additional categorization of these events involves proximity events where above 90% of the defined minima are maintained and an operational error is a violation below the 90% threshold. Figure 19 below presents the total counts of this breakdown. Of interest here is that the most consistently difficult boundary change condition-High magnitude change- was the only one to not experience any operational errors. The Medium change condition had by far the most. However, of all of the actual operational errors, an investigation into each revealed extenuating circumstances that either related to simulation artifacts or had greater implications for future concepts related to communications and trajectory based operations. None of the cases merited the drawing of any conclusions of safety based explicitly on any of the boundary change conditions or the DAC concepts as a whole.

Boundary Change Condition	Proximity Events	Operational Error (LOS)
Baseline	2	1
Low Magnitude	1	2
Medium Magnitude	0	5
High Magnitude	0	0

Figure 19. Total counts of Proximity events and operational errors defined by a loss of separation.

F. Acceptability

Following each run, participants were given a questionnaire created specifically for that run. A number of questions were presented in the questionnaire as well as an image of each boundary change as a reminder. One of the questions asked for a rating, on a scale of one to seven with seven being the strongest, concerning the acceptability of each boundary change. While each boundary change was different, the results presented here take an overall look at acceptability as it applies to the boundary change conditions rather than specific boundary configurations. Descriptive statistics for acceptability reveal a downward trend (see Figure 20) in acceptability as the magnitude of change increased. The Low magnitude change condition resulted in the highest mean acceptability rating with 6.66 ($SD= 0.37$) while the Medium change condition was slightly lower with a mean rating of 6.04 ($SD= 0.82$). The High magnitude change condition had a much lower acceptability rating for its boundary changes with a mean rating of 4.46 ($SD= 1.12$). A one way, repeated measures ANOVA was conducted to examine the differences

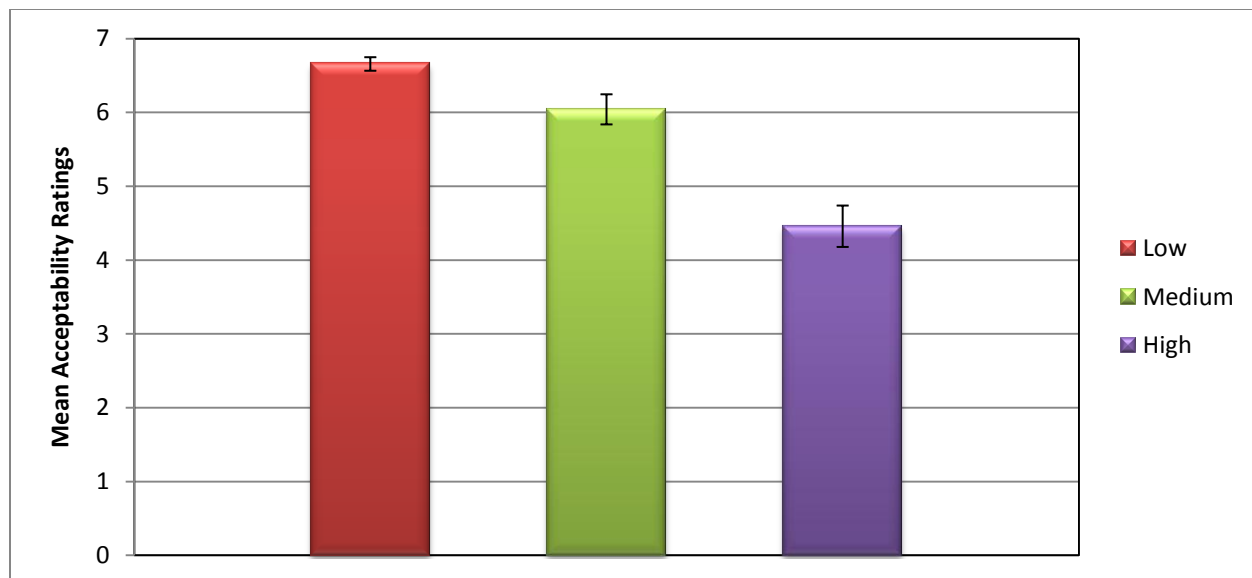


Figure 20. Mean post-run acceptability ratings for boundary configurations experienced.

in the reported mean acceptability ratings where a significant difference was found, $F(2, 45) = 30.00$, $p < .001$. A Bonferroni post hoc test revealed significant differences between the Low and High magnitude change conditions as well as the Medium and High change conditions meaning that the High magnitude change condition's boundary configurations were significantly less acceptable to the participants than either the Low or Medium change boundaries.

V. Discussion

This was the first in a planned series of human-in-the-loop studies into the Dynamic Airspace Configuration concept. The answers sought through this study were for questions concerning some of the basic, fundamental issues surrounding the concept and its potential impact on the airspace system. The results presented in this paper provide insight into the different forms of impact that some of the different algorithmic approaches had both on traffic and the human operators involved in this exploration.

Three different DAC algorithms were initially selected for their different approaches to the DAC concept as well as for their relative differences in terms of aggressiveness in the changes to sector geometry. The resulting sectorizations from the three algorithms were categorized for this study into Low, Medium, and High magnitude change. The first analyses conducted as part of the larger effort concerned the actual airspace changes that took place in the final sectorizations used in the study with an eye toward quantifying the differences between them. This was done through a comparison of the percentages of airspace gained and lost for each boundary change relative to the previous boundary. In support of the categorization of the sectorizations into the three levels, the results for airspace gained showed a significant difference between the Low and High magnitude change sectorizations and a strong yet non-significant difference between the Low and Medium change conditions. Results for airspace lost showed even stronger support for the categorization where significant differences were found between each of the three levels of boundary change magnitude. Confirming these categorizations was a necessary first step in the analysis due to the fact that this would be the context through which the other results would be framed and was also an integral aspect of how the study was designed and conducted.

The next step in the analysis was to look at how each of the boundary change conditions, this time including the Baseline condition, managed to effectively distribute traffic loads between sectors. This is, after all, one of the primary purposes of the DAC concept and is therefore important to see how the different approaches to the problem handle this task. The results reported in the previous section for this concern were descriptive in nature where the standard deviation of sector occupancy counts between the four test sectors was plotted over time for each of the four boundary change conditions. Not surprisingly, Baseline showed the greatest amount of variability due to the fact that this was intentional through the construction of the traffic scenarios. Results at this juncture

basically show that the Baseline condition was a decent representative example in support of the need for dynamic resectorization to address demand-capacity imbalances and would be an adequate comparison point. The patterns for the three other boundary change conditions, however, were not as straightforward as the Baseline, where some approaches performed better than others depending on the characteristics of the traffic peaks that they were in response to. The traffic scenarios had two peaks in the traffic pattern with the first exceeding and sustaining traffic counts above the MAP value before dipping back below. The duration of this first peak was not as sustained as that of the later peak, however, and it appears as though this made somewhat of a difference in how the load imbalance was handled. In Figure 11, one can see that for the shorter peak, the Low magnitude change condition managed to better distribute the traffic followed by the High magnitude change condition and then the Medium change condition. However, as the scenarios progress to the longer peak, there is a reversal where the High magnitude change condition appears to more successfully manage the traffic distribution with the Medium change condition showing a similar behavior. It is interesting to note that the behavior of the traffic distribution changed in relative terms between the different approaches depending on the traffic characteristics. This suggests a potential avenue for further investigation into the relationship between the different sectorizations and particular types of traffic patterns that they might handle. With respect to the handling of the load imbalances during the two scripted peaks, it appears as though the High magnitude change condition did the better job but, as the other results showed, it was at the expense of other factors.

Another indicator of effective traffic distribution was addressed through the analysis of duration of aircraft count over the established MAP value of 22. The trend observed through the analysis was downward meaning that the Baseline condition sustained the greatest number of minutes that the test sectors had aircraft counts in excess of MAP with each successive increase in boundary change aggressiveness resulting in fewer numbers of minutes. This meant that in the end, the condition with the shortest duration of MAP excess was the High magnitude change condition.

Despite the results just discussed showing the effectiveness of the High magnitude change condition, there were other areas that pointed to potential problems. One such case was observed through an analysis of aircraft dwell times, which refers to how long aircraft spent in a sector. For these results, shorter time durations can be interpreted as being less desirable because what works to bring duration times down are transit durations that are short. The factors that can produce these short durations are sector characteristics such as narrow aspect ratios and boundary characteristics like jagged edges as observed in the High change sectorizations. In the end what was observed was that the High magnitude change condition had a significantly shorter duration for dwell time compared to that of the Baseline condition, which also means that an increased level of taskload and workload would likely accompany shorter dwell times due to the extra coordination that would be required in response.

The number of handoffs and point outs across the four boundary change conditions was examined as a partial reflection of taskload and, as just referenced, an inverse relationship between dwell time and taskload was observed. More specifically, for both the comparisons of handoffs and point outs, each progressive level of boundary change magnitude resulted in an increased number of both tasks such that the High magnitude condition had the greatest numbers of both. Analyses of each task showed a significant difference in means for handoffs with a marginally significant difference between the Baseline and High magnitude change conditions. No significant difference was found with point outs, but the increasing trends for both tasks are clear and consistent and seem to correlate with dwell time. One note about the results for point outs is that the numbers obtained were the minimum and do not provide a full accounting of them. In busy and complex times, a number of point outs were made verbally and for groups of aircraft at a time when certain sector configurations affected many aircraft. It could be the case that if there were a full accounting, significant differences would be observed as they are for handoffs.

Just as the numbers of handoffs and point outs increased along with boundary change magnitude, so too did the reported workload both in terms of overall workload and the ratings that surrounded boundary changes only. While the tasks surely contributed to workload it is just as certain that there were other factors worth investigating that contributed to workload as well. The results presented in Figure 16 show overall mean workload ratings that increased with boundary change condition. There was also another piece of data in this presentation that is worth noting, however, and that is the trend in the standard deviation. As the workload increased for each change condition, the standard deviation values decreased. This follows the earlier results for traffic distribution where counts were distributed more evenly among the sectors. As more traffic was evenly distributed, so too was the workload such that in the High magnitude change condition the participants were almost uniformly reporting the same level of high workload. These trends remained constant for workload reported at the boundary changes with a slight decrease in the Low change condition, a slight increase in the Medium change condition, and the greatest increase in the High change condition. The decrease in the Low change condition can be interpreted as meaning

that other factors unrelated to the boundary changes resulted in higher workload that was removed once the ratings were isolated to the boundary change points. The opposite applies for the Medium and High change conditions where the times of lower workload tended to mask the increased levels of workload associated with the boundary changes.

While workload and taskload do have implications for safety, the results did not entirely follow the same trend nor suggest that any one type of boundary change condition or the DAC concept as a whole was any less safe than the Baseline. In terms of the number of conflicts, there was no significant difference in the number between any of the boundary change conditions. This speaks to one of the virtues of this concept due to the fact that the changing of aircraft trajectories is unnecessary to manage the airspace and demand-capacity imbalance. With more trajectory changes comes greater complexity and the greater possibility of creating conflicts that could result in losses of separation. One aspect of safety that did follow the earlier trends was in the area of operational errors with respect to airspace deviations. For the three components that were analyzed (late handoff, late handoff bypassing owned sector, and late acceptance of handoff), each of the boundary change condition in order of change magnitude showed an increasing number of each of the deviations where the High change condition had the greatest overall numbers for each of the components. A comparison of the means showed that the High change condition had significantly higher numbers of deviations which could have an impact on safety given the right conditions. However, a look at the number of proximity events and separation violations does not support this idea. Looking at the number of actual losses of separation shows that the High change condition actually had none whereas Baseline had one occurrence, the Low change condition had two, and the Medium change condition had five. An examination of each of these occurrences with respect to the actual events that led to the violation did not reveal any consistencies that could be attributed to anything related to the DAC concept as a whole or any of the particular boundary change conditions.

The final metric that was analyzed was the acceptability of the boundary changes in each of the boundary change conditions (without Baseline). This was an important metric to include because the acceptability ratings share a relationship with feasibility in terms of the concept and the ability of the controllers to work in the type of environment envisioned in the concept. As seen in Figure 20, there is a decreasing trend in the level of acceptability where the Low change condition had the highest acceptability rating followed by the Medium change condition with slightly less, and a more marked decline in acceptability for the boundaries presented in the High change condition.

While analyses of the data from this study are ongoing, the results presented thus far provide some direction on avenues to take in subsequent research and development. In some ways the results were paradoxical in that as the magnitude of boundary change increased, the management and distribution of traffic became more effective. However, this came with the added cost of taskload and workload and ultimately had a negative impact on acceptability. This suggests that there is not a singular solution or way forward but that, at least with the approaches implemented in this study, there are benefits and useful components in each approach that are worth investigating and considering further. It is also likely that these different benefits are realized through different situations and traffic environments such that a more inclusive or patchwork approach may be the most effective way forward.

Feedback from the participants regarding factors that they felt impacted their ability to satisfactorily control traffic and negatively impacted their acceptability of particular boundary changes provide further guidance on how to move forward both in terms of algorithm and concept development. The factors that they cited as important were large changes in airspace volume, large modifications to the shape of the boundary, large changes in the number or orientation of adjacent sectors and, similarly, changes in the sequence/position of upstream or downstream sectors, large changes in traffic flow, high frequency of boundary changes, and not having adequate preview time in which to handle the transition. Other feedback from the participants pointed to what they felt were key enabling technologies for such a concept as DAC to reasonably be put in place at least in terms of what they experienced in the instantiation of the concept in this study. These technologies were almost all agreed upon, which were data link communications with automatic transfer of control, conflict probing, and having the support of automated conflict resolution.

Taken together, the results and feedback from the study showed that DAC is a promising concept worth further development and refinement. The results also present researchers with a number of items and issues with which to consider and incorporate into their work in making DAC viable. It appears as though a number of tradeoffs are required if some middle ground is to be met between how the demand-capacity imbalance is most effectively addressed and how best to do that while keeping the human controller integrated and functioning meaningfully within the system. Based on the results from this study, further research can begin in addressing these issues.

Acknowledgments

This work was funded by NASA's Airspace Systems Program and NGATS-ATM Airspace Project. The authors would like to thank the participants in this study for their hard work and unending patience. Thanks must also go to the MACS development team for making the study possible.

References

- Brinton, C. and Pledgie, S. (2008) Airspace Partitioning using Flight Clustering and Computational Geometry. In Proceedings of the 27th Digital Avionics Systems Conference (DASC), St. Paul, MN.
- Corker, K., Liang, D., Lee, P., & Prevôt T. (2007) New Air Traffic Management Concepts Analysis Methodology: Application to Multi-Sector Planned in US Airspace. *Air Traffic Control Quarterly*. Vol. 15, no. 4, pp. 347-367.
- Drew, M. (2008). Analysis of an Optimal Sector Design Method. In Proceedings of the 27th Digital Avionics Systems Conference (DASC), St. Paul, MN.
- Erzberger, H. (2001). *The Automated Airspace Concept*. 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, NM, USA.
- Federal Aviation Administration (2009). FAA Aerospace Forecast Fiscal Years 2009-2025. Washington, DC, p.31.
- Hoffman, R., Prete, J. (2008). "Principles of Airspace Tube Design for Dynamic Airspace Configuration," AIAA-ATIO Conference, Anchorage Alaska.
- Klein, A., Rogers, M., and Kaing, H. (2008). Dynamic FPAs: A New Method for Dynamic Airspace Configuration. Integrated Communications Navigation and Surveillance (ICNS) Conference. Bethesda, MD.
- Klein, A., Rogers, M., and Leiden, K. (2008). Simplified Dynamic Density: A Metric for Dynamic Airspace Configuration and NEXTGEN Analysis. In Proceedings of the 27th Digital Avionics Systems Conference (DASC), St. Paul, MN.
- Kopardekar, P., Bilimoria, K., & Sridhar, B. (2007). Initial concepts for Dynamic Airspace Configuration, 7th Aviation Technology, Integration and Operations (ATIO) Seminar. AIAA, Belfast, Northern Ireland.
- Kopardekar, P., & Magyarits, S. (2003). Measurement and prediction of dynamic density. 5th USA/Europe Air Traffic Management R&D Seminar, Budapest, Hungary, June, 2003.
- Kopardekar, P., Schwartz, A., Magyarits, S., & Rhodes, J. (2007). Airspace complexity measurement: an air traffic control simulation analysis. 7th USA/Europe Air Traffic Management R&D Seminar, Barcelona, Spain, June, 2007.
- Lee P., Corker K., Smith N., Prevôt T., Martin L., Mercer, J., Homola J., & Guneratne E. (2006). A Human-in-the-loop Evaluation of Two Multi-Sector Planner Concepts: Multi-D and Area Flow Manager. In Proceedings of the 25th Digital Avionics Systems Conference (DASC), Portland, OR.
- Lee, P.U., Mercer, J., Gore, B., Smith, N., Lee, K., & Hoffman, R. (2008). Examining Airspace Structural Components and Configuration Practices for Dynamic Airspace Configuration, AIAA Guidance, Navigation, and Control Conference and Exhibit 18 - 21 August 2008, Honolulu, HI.
- Prevôt, T. (2002) Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi-Aircraft Control System MACS. In S. Chatty, J. Hansmann, & G. Boy. (Eds). *HCI-Aero 2002*, AIAA Press, Menlo Park, CA. pp 149-154.
- Programme for Harmonised Air Traffic Management Research in Eurocontrol (PHARE) (1997). En-route Multi Sector Planning Procedures. PHARE document DOC 97-70-15.
- Super Highway (2008). D3.1 Conclusions Report. SPH-WP3-v0.4-ISD-059-V1.0-ED-CO
- Zelinsky, S. (2009). A Comparison of Algorithm Generated Sectorizations. Eight USA/Europe Air Traffic Management Research and Development Seminar (ATM 2009), Napa, CA.