# Evaluation of NextGen Air Traffic Control Operations with Automated Separation Assurance

**Thomas Prevot, Joey Mercer, Lynne Martin, Jeffrey Homola, Christopher Cabrall, Connie Brasil**
San Jose State University / NASA Ames Research Center
Moffett Field, CA 94035, USA
thomas.prevot@nasa.gov

## ABSTRACT

As part of an ongoing research effort into functional allocation in a NextGen environment, a controller-in-the-loop study on ground-based automated separation assurance was conducted at NASA Ames' Airspace Operations Laboratory in February 2010. Participants included six FAA front line managers, who are currently certified professional controllers and four recently retired controllers. Traffic scenarios were 15 and 30 minutes long where controllers interacted with advanced technologies for ground-based separation assurance, weather avoidance, and arrival metering. The automation managed the separation by resolving conflicts automatically and involved controllers only by exception, e.g., when the automated resolution would have been outside preset limits.

Results from data analyses show that workload was low despite high levels of traffic, Operational Errors did occur but were closely tied to local complexity, and safety acceptability ratings varied with traffic levels. Positive feedback was elicited for the overall concept with discussion on the proper allocation of functions and trust in automation.

### Keywords
Separation assurance, simulation, NextGen

## INTRODUCTION
The Joint Planning and Development Office (JPDO) has identified the action area "Air/Ground Functional Allocation" as a high priority [1]. Its concern is to address the "lack of clarity in the allocation of new functions to the aircraft and flight crew (includes human/automation as well as avionics/ground automation allocations)." This paper presents results from a human in the loop (HITL) study in the Airspace Operations Laboratory at NASA Ames Research Center that examined the functional allocation between air traffic controllers and automation within the concept of ground-based automated separation assurance. In a separate publication [2] this ground-based approach to separation management is compared with the approach of airborne separation management investigated at NASA Langley Research Center.

In this paper, we first discuss the primary problem of safely doubling or tripling airspace capacity in the next two decades. Next, we describe the approach of allocating many separation assurance functions to the ground-based automation. This approach was initially investigated in a sequence of part-task studies before the most recent experiment simulated the operations in a more comprehensive air traffic control environment. After presenting critical elements of this method a set of initial findings related to acceptability, safety and workload will be discussed.

## PROBLEM
The JPDO expects NextGen to "accommodate significantly increased traffic levels with broader aircraft performance envelopes and more operators within the same airspace, increasing the complexity and coordination requirements when air traffic management (ATM) is required" [3]. The required capacity increase ranges from 1.4 times to 3 times the baseline capacity of 2004.

The main factor limiting en route capacity is controller workload associated with providing safe separation between aircraft. In today's very safe system, air traffic controllers take active control over each aircraft in their airspace and issue clearances to keep them separate from other traffic, expedite traffic flows, and provide additional services, workload permitting. Being actively involved with each flight provides the awareness required to detect and resolve potential losses of separation, even without advanced automated aids. This manual process, however, can only be performed for a limited number of aircraft. In recognition of this fact, each airspace sector today has a defined maximum number of aircraft that are allowed to enter. This constraint exists as a way of ensuring that the demands on the cognitive resources of the air traffic controller(s) controlling this sector are not exceeded.

Assuming that this level represents the sustained traffic load a controller can comfortably manage today, a fundamental change in operations has to occur to meet the projected traffic levels for NextGen. Figure 1 shows a current day display at the envisioned traffic levels and an experimental display designed for automated separation assurance operations.

For NextGen, it is envisioned that trajectory-based operations will replace clearance-based operations in many parts of the airspace. New automated separation assurance functions are intended to help overcome the aforementioned limitations of controllers in manually maintaining safe separation between aircraft. The two primary separation assurance (SA) concepts are ground-based automated separation assurance [4] and airborne separation management [5]. The current approach to ground-based automated separation assurance was informed by prior HITL simulations [6, 7] and is presented next.

### Approach: Ground-Based Automated Separation Assurance
The general idea of ground-based automated separation assurance is to let the automation monitor and/or manage nominal trajectory-based operations of equipped aircraft (low-lighted on the bottom display in Figure 1), while the operator handles off-nominal operations, provides additional services and makes decisions on situations that are presented to

**Figure 1: NextGen traffic on current day controller display (top) and experimental display (bottom)**

him/her (high-lighted on bottom display in Figure 1). The separation responsibility resides with the service provider, which means the air traffic controller and ground-based automation. The primary difference from today's system is that the ground-based automation is responsible for conflict detection and that separation assurance automation generates conflict resolution trajectories integrated with data link. The modified trajectories are sent to the aircraft either by the controller or directly by the ground-based automation, whenever certain predefined criteria are met. The flight crews' responsibilities related to separation assurance do not change from today.

## Concept of Operations

*NextGen Air Traffic Environment*

The concept of automated separation assurance is enabled by integrating controller workstations, ground-based automation, data link, flight management automation, and flight deck interfaces. The ground automation creates, maintains, and communicates trajectories for each flight. The air traffic environment is generally based upon the mid-term environment for the high altitude airspace outlined by the Federal Aviation Administration (FAA) [8] and assumes the following characteristics: Each aircraft entering high altitude airspace is equipped with an FMS that meets a required navigation performance value of 1.0 and have integrated data link for route modifications, frequency changes, cruise altitudes, climb, cruise, and descent speeds, similar to those available in current day Future Air Navigation System 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/ descent speeds, and weight are communicated to the ATC system. The goal is to make conflict detection highly reliable and to detect trajectory-based conflicts with enough time before initial Loss Of Separation (LOS). However, some sources of trajectory uncertainties remain and include 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 the 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 loss of separation. When conflicts are detected within a short time before LOS, an automated tactical conflict avoidance function can generate heading changes and send this information to the flight deck via a separate high-priority data link path.

## New Technologies for Air Traffic Controllers

In order to provide the required automation support to the controller, a new NextGen ATC workstation prototype was developed built upon an emulation of the operational enroute controller system. The workstation provides access to key functions that support the operator in managing high traffic densities effectively. Figure 2 shows the controller display as implemented in NASA's Multi Aircraft Control System (MACS) the software that is used for this research [9].

All functions for conflict detection and resolution, trajectory planning and routine operations are directly accessible from the tactical controller display. Transfer of control and communication between controllers is conducted by the automation close to the sector boundaries. Nominally, aircraft are displayed as chevrons with altitudes, a design originally developed for cockpit displays of traffic information [10]. Traffic conflict information, hazard penetration and metering information is presented where applicable. Full data tags are only displayed in short-term conflict situations, or when the controller selects them manually. Time-based metering is supported via timelines and meter lists. The timelines show aircraft's estimated and scheduled arrival times at specific fixes, which are often meter fixes into congested airports.

The controller can request trajectories to avoid traffic conflicts, weather hazards and solve metering conflicts via various easy-to use mechanisms: via use of keyboard entries, data tag items, the conflict list and/or the timeline. The automated trajectory-based conflict resolutions are generated by an autoresolver module originally developed as part of the Advanced Airspace Concept [4]. When initiated by the controller, the automatically generated trajectory becomes a trial trajectory (indicated in cyan in Figure 2). All trajectory changes are immediately probed for conflicts and provide

real-time feedback on their status, before they are sent. Therefore, the tools are designed to be interactively used. The controller can then modify and/or uplink the trajectory constraints to the aircraft.
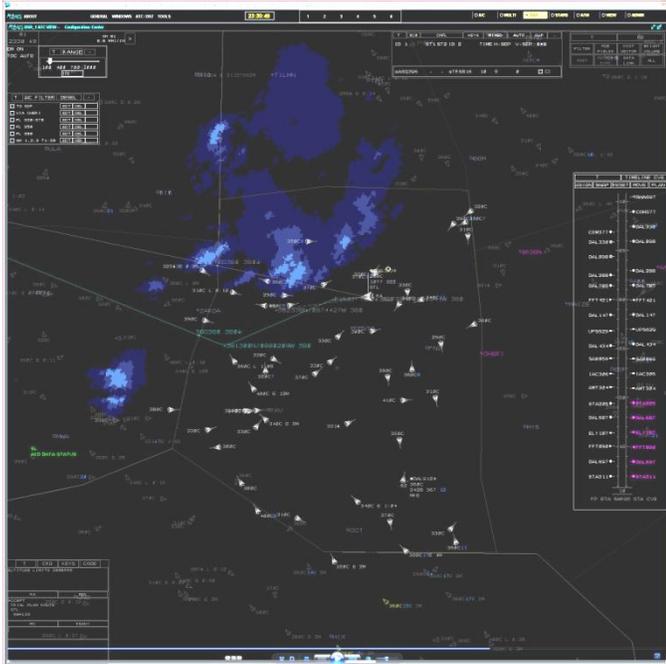


**Figure 2: Controller display at NextGen traffic levels with weather and metering**

As mentioned before, the automation highlights short-term conflicts and computes tactical heading changes for one or both of the involved aircraft within three minutes of predicted LOS. These resolutions are generated by a Tactical Safety Enhanced Flight Environment (TSAFE) module [11]. In the current study, the automation sent the heading change(s) at two minutes to LOS and the controller had no means to intervene. This was an experimental configuration decision rather than an operational concept decision.

## METHOD: Experimental Design

The experiment was designed to meet two objectives:

A. Analyze air traffic control operations with ground-based automated separation assurance at different NextGen traffic levels under typical constraints, such as arrival time constraints and convective weather

B. Compare this ground-based automated separation management approach to the airborne approach investigated at NASA Langley's Air Traffic Operations Laboratory (ATOL).

This paper will discuss results of the analysis for objective A; the comparison for Objective B will be presented in a separate publication [2]. The traffic scenarios and scenario events were tightly controlled. They included *short* scenarios of 15 minutes length and *medium* scenarios of 30 minutes length. *Long* scenarios of 3 hours length were also run, but have not yet been analyzed.

### Traffic level and arrival time constraints

The *medium length scenarios* were intended to investigate the effect of traffic level and arrival time assignments. Therefore

they were designed as a 2x2 repeated measures matrix that varied the presence/absence of arrival time constraints over two NextGen (Traffic) Levels within subjects.

NextGen Level A represents an increase of approximately 50% of aircraft operations across the entire test airspace and meets the airspace operations forecast for 2025 as published by the FAA in 2009 [12]. Level A results in a wide spread of throughput levels for individual sectors ranging from current day values (15 aircraft) to more than twice current day values (42 aircraft). NextGen Level B represents an increase of approximately 100% of aircraft operations across the entire test airspace and is introduced to test scalability. Level B results in individual sector throughput of up to three times current day values (55 aircraft).

| **Arrival Time Constraints** | No (Baseline) | M1 | M2 |
|---|---|---|---|
| | Yes (STA) | M4 | M3 |
| | | A | B |

**NextGen (Traffic) Level**

**Figure 3: Design matrix for 30-minute (*medium*) scenarios**

The *short scenarios* were designed to investigate the impact of synchronous vs. dispersed trajectory changes on the operations. To force trajectory changes, arrival times were re-assigned during the runs, prompting the operators/automation to generate new arrival trajectories. The timing of arrival time changes was varied within subjects. All the short scenarios were ran with the NextGen Level B traffic density..

| **Timing of Arrival Time changes** | Basic (STA) | S1 |
|---|---|---|
| | Dispersed (one every minute) | S2 |
| | Synchronous (all at 6 or 8 minutes) | S3 |

**Figure 4: Design matrix for 15-minute (*short*) scenarios**

### Test Airspace

The primary test airspace is located in the central United States and spans the eastern part of Kansas City Center (ZKC) as well as the western part of Indianapolis Center (ZID).
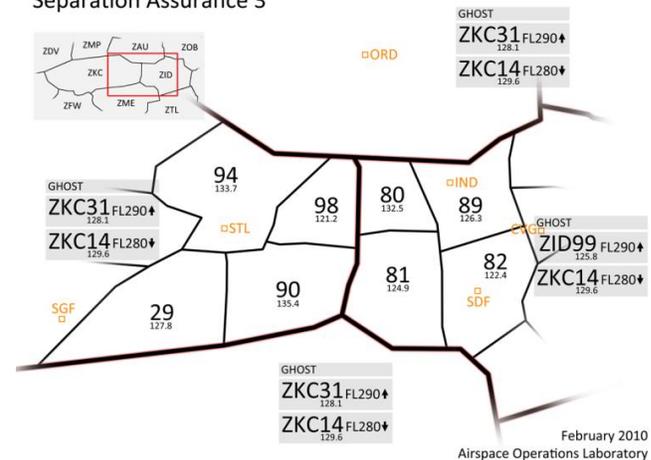


**Figure 5: Test airspace. Short and medium scenarios used sectors ZKC90, ZKC98, ZID80 and ZID81.**

As depicted in Figure 5, it covers high altitude airspace at flight level 290 and above. In the short and medium scenarios, the focus was on the four central sectors (ZKC90, ZKC98, ZID80 and ZID81), which were staffed by participant controllers.

### Laboratory

The simulation was conducted in the Airspace Operations Laboratory (AOL) at NASA Ames Research Center, which uses the MACS as its simulation and rapid prototyping software [13]. For this study the AOL was configured with two participant control rooms, each hosting four tactical air traffic control sector positions and one supervisor position.

Four "ghost" confederate controllers managed traffic flows in and out of the test sectors and ten general aviation pilots operated the simulated aircraft throughout the test airspace. During the study runs, the two control rooms were run independently in

**Figure 6: Air traffic control room in the AOL during study**

separate "worlds" each with their own ghost controllers and simulation pilots.

### Participants

Participant positions were the sector positions and the supervisor positions. These were staffed with six front line managers from six different en route facilities, who were current on the radar position. For the short and medium runs four additional recently retired en route controllers staffed the remaining positions within the test area. The FAA participants had not been exposed to this concept and –except for one– any of the technologies before. During the Medium runs the FAA participants staffed sectors 90, 98, and 81. During the Short runs they rotated through sectors 90, 98, and the supervisor position. The retired controllers helped during prior system shakedowns and simulations and were familiar with the tools and concepts before the simulation started.

### Data collection

The study was conducted during eight days in February/March 2010. During the first morning, all participants were briefed, and then the retired controllers gave a demonstration of the operations to the FAA participants. The FAA participants were trained for 2.5 days. Data collection was conducted during days 4, 5, 7, and 8. The short and medium runs were conducted in the morning. During the study the comprehensive built-in MACS data collection system recorded all possibly relevant data, including all aircraft states, trajectories, automation, and operator events. All displays were recorded by a commercial screen capture product. At the end of each run participants responded to a short questionnaire. After the study, participants completed a long questionnaire that included items on function allocation and trust in the automation. All questionnaires (post-run and post-

simulation) were posted electronically at the participant position.

## RESULTS

The results presented in this paper include the aircraft counts within the sectors, the controller workload, operational errors and qualitative feedback on acceptability of the operations and the functional allocation between controller and automation.

### Aircraft counts

*Medium Duration Runs*

Figure 7 presents the aircraft counts, per test sector, plotted over the duration of the medium runs (30 minutes). On the left hand side of the figure is the Level A condition, which was the lower overall count of the two, where it can be seen that sector ZKC90 and ZID81 –the largest of the four sectors– had the highest counts in order, followed by ZKC98 and ZID80–the smallest of the four sectors.

In terms of aircraft count, the mean number of aircraft occupying ZKC90 for the Level A condition was 34.69 ($SD$ = 4.51), followed by ZID81 with a mean number of 25.94 aircraft ($SD$ = 5.29) in sector. Sector ZKC 98 had a mean number of 20.86 aircraft ($SD$ = 2.20), followed by sector ZID80, which had the lowest aircraft count with a mean number of 13.91 aircraft ($SD$ = 3.71). In addition to these numbers, it can be seen from Figure 7 that there were negligible differences in aircraft count between the Baseline and STA conditions.
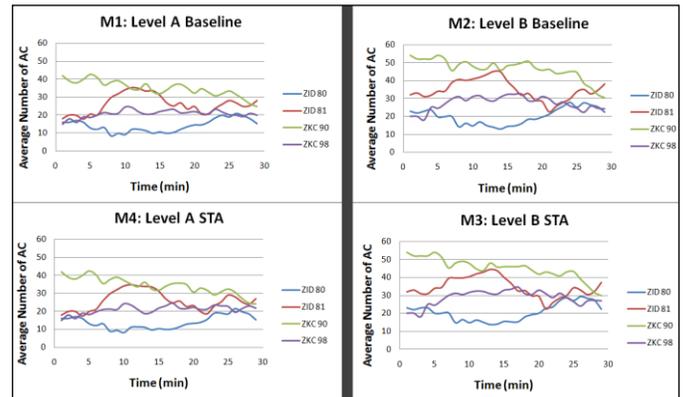


**Figure 7: Average traffic count over time for the Medium duration condition**

The Level B condition, found on the right hand side of Figure 7, had higher levels of aircraft count in each of the sectors relative to that experienced in Level A. While the actual counts differed between the conditions, the order of sectors in terms of aircraft count held constant across the Level B conditions with ZKC90 clearly having the highest counts ($M$ = 45.76, $SD$ = 5.90), followed by ZID81 ($M$ = 34.66, $SD$ = 5.62), ZKC98 ($M$ = 28.00, $SD$ = 4.10), and ZID80 ($M$ = 20.34, $SD$ = 4.75). And just as in the Level A condition, aircraft count did not differ significantly between the Baseline and STA conditions.

*Short Duration Runs*

As mentioned earlier, the Short duration runs used NextGen traffic level B for all three scheduling conditions (S1, S2, and S3). The differences in terms of the scheduling task between the three conditions did not translate into differences in

aircraft count with each plot mirroring the other. The relative order of aircraft count by sector is the same as the Medium conditions with ZKC90 experiencing the highest counts ($M = 49.40$, $SD = 3.24$), followed by ZID81 ($M = 37.75$, $SD = 4.96$), ZKC98 ($M = 26.77$, $SD = 4.53$), and finally ZID80 with the lowest aircraft count of the four sectors ($M = 18.45$, $SD = 3.63$).

**Workload**

Subjective workload was obtained during the simulation runs through a workload assessment keypad. Workload prompts were presented every three minutes throughout the run to which the participants rated their perceived workload on a scale from one to six with six being the highest level of workload possible.

*Medium Duration Runs*

In the Medium runs, the conditions were Baseline without any scheduling tasks and STA with scheduling across the Level A and B traffic levels. Figure 8 presents the mean workload ratings for each of these conditions.
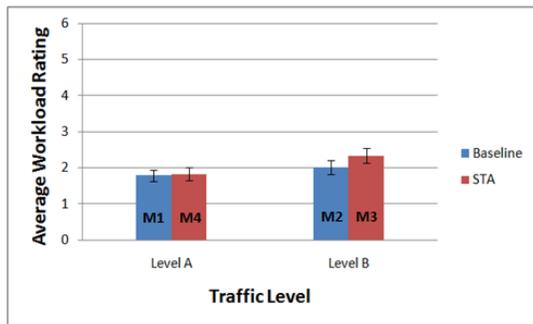


**Figure 8: Mean reported workload for the Medium duration condition.**

Descriptive statistics for these conditions show that for Baseline, Level A (M1), the mean workload rating was 1.78 ($SD = 0.64$) and the corresponding STA condition (M4) had a slightly higher mean rating of 1.83 ($SD = 0.74$). At the Level B traffic level, the Baseline condition (M2) mean workload rating of 2.01 ($SD = 0.79$) was higher in comparison with the Level A results. The STA condition (M3) had a more noticeable increase relative to the Baseline condition with a mean rating of 2.34 ($SD = 0.79$). Overall, the STA conditions had slightly higher workload ratings, although even with the increase the ratings denote low levels of workload.

To investigate the described differences, a one-way repeated measures Analysis of Variance (ANOVA) was conducted to test the difference between the Baseline and STA conditions. Despite the trend for the STA condition having higher mean workload there were no significant differences between it and Baseline ($F(1, 31)= 2.60$, $p> .05$). Workload across the A and B traffic levels was examined next where the differences appeared stronger, with Level B producing significantly greater workload than Level A ($F(1, 31)= 14.85$, $p< .01$).

To provide a more complete picture of workload, additional analyses were conducted to examine how workload differed between the test sectors. Similar to the differences in aircraft count between the sectors, Figure 9 shows how there were differences in mean workload reported throughout the runs.

As observed in Figure 9, the sector with the most consistently highest reported workload was ZKC90 (shown in green). Descriptive statistics show that this was indeed the case with a mean reported rating of 2.57 ($SD = 0.99$). ZKC98 (in purple) was next with a mean workload rating of 2.09 ($SD = 0.58$) followed by ZID81 (in red) ($M = 1.80$, $SD = 0.57$), and ZID80 (in blue) with the lowest mean reported workload ($M = 1.49$, $SD = 0.34$).



**Figure 9: Mean workload ratings reported over time, per sector, for the Medium duration condition.**

A one-way ANOVA was conducted on the mean workload ratings between the four sectors to further examine the differences. A significant difference between sectors was found ($F(3, 60)= 7.65$, $p< .01$). To further examine the differences between sectors, Tukey's Honestly Significant Difference (HSD) tests were conducted where it was found that ZKC90 had significantly higher workload than ZID81 ($p< .01$) and ZID80 ($p <.01$). The only sector that ZKC90 did not significantly differ from was ZKC98.

In summary, analyses of workload in the Medium condition did not result in significant differences between the Baseline and STA conditions, but significant differences were found between traffic level conditions with Level B resulting in higher levels of workload than in A. At the sector level there were significant differences between sectors with ZKC90 reporting significantly higher levels of workload relative to all sectors with the exception of ZKC98. Despite these differences, the mean workload ratings were still fairly low considering the levels of traffic experienced.

*Short Duration Runs*

Workload was also analyzed for the Short runs across the three scheduling conditions –basic scheduling (S1), Dispersed STA updates (S2), and Synchronous STA update (S3).

Figure 10 presents the mean workload ratings for each of the three conditions where a positive trend can be observed starting with the lowest mean workload being reported in the S1 condition ($M = 1.97$, $SD = 0.89$), followed with slightly higher mean workload reported in the S2 condition ($M = 2.10$, $SD = 0.97$), and the highest reported ratings in S3 ($M = 2.32$, $SD = 1.11$). As with the Medium runs, even the highest mean workload ratings were still quite low. To examine the observed differences between conditions in the Short runs, a one-way repeated measures ANOVA was conducted where a

significant difference was found ($F(2, 94)= 4.81$, $p< .05$). Tukey's HSD tests were conducted to further examine the differences where it was found that the S3 condition had significantly higher workload than the S1 condition ($p< .01$).
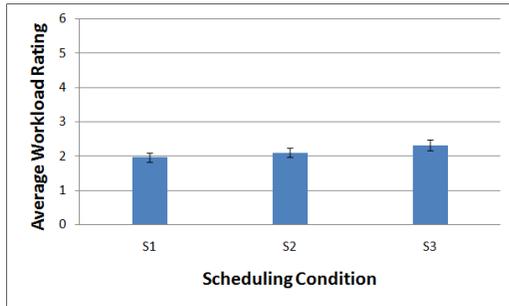


**Figure 10: Mean reported workload for the Short duration condition.**

Workload was further examined to take into account possible differences between the four test sectors. Figure 11 presents the mean workload ratings for each of the sectors plotted over the 15 minute duration of the run. Here it can be seen that similar to the Medium runs, sector ZKC90 had, comparatively, a high level of workload. However, the workload reported from participants working ZKC98 show very similar workload ratings to ZKC90. Descriptive statistics for the sectors reveal that ZKC90 reported the highest mean workload ratings, much like in the Medium duration runs, ($M = 2.88$, $SD = 0.86$), followed closely by ZKC98 ($M = 2.81$, $SD = 0.76$), then by ZID81 ($M = 1.51$, $SD = 0.71$), and ZID80 ($M = 1.32$, $SD = 0.34$).

To further analyze these workload differences between the four test sectors, a one-way ANOVA was conducted. Just as in the Medium condition, a significant difference was found ($F(3, 140)= 50.94$, $p= .00$). Following this result, Tukey's HSD tests were conducted where it was found that ZKC90, again, had significantly higher workload than ZID81 and ZID80 ($p< .01$), but did not differ significantly from ZKC98. Unlike in the Medium condition, however, ZKC98 also differed significantly from ZID81 and ZID80 ($p< .01$) with higher mean workload.

To summarize, results from the Short condition workload analyses revealed significant differences between the Synchronous and basic scheduling conditions. Significance was also discovered while examining workload at the sector level with Kansas City sectors reporting significantly higher
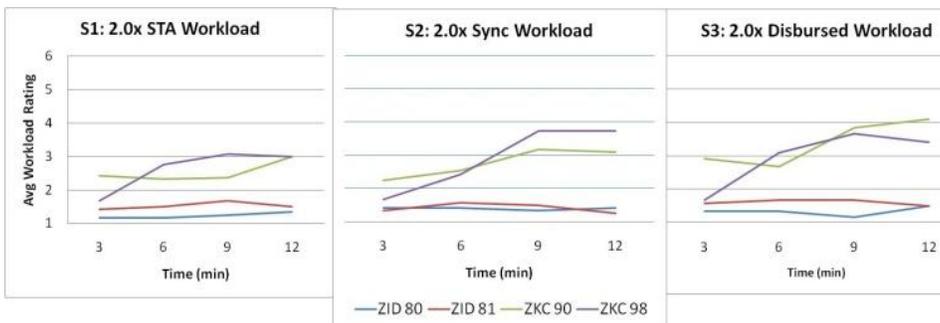
levels of workload than Indianapolis sectors.

## Loss of Separation Analysis

Certain traffic flow interactions of climbing and descending aircraft within the scenarios included extremely complex elements that stressed the functional allocation concepts, provoking hard-to-resolve short-term conflicts in certain areas. These sometimes resulted in losses of separation that were categorized into Operational Errors (OE) (maneuvers with less than 4.5 nm and 800 feet spacing and proximity events (between 4.5 and 5 nm). The OEs were analyzed with univariate ANOVAs (see Figure 12).
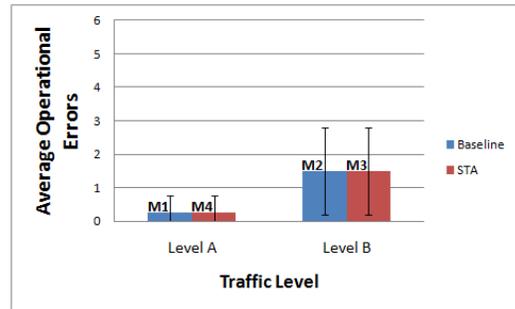


**Figure 12: Average number of operational errors during medium runs.**

The medium runs gave insight into the effects of traffic levels and the arrival time metering on the safety of the operations. A two-way univariate ANOVA found a significant effect of traffic level on the number of OEs such that there were significantly more in the Level B ($M = 1.50$, $SD = 1.19$) versus the Level A ($M= 0.25$, $SD = 0.46$) condition ($F(1,12) = 6.52$, $p < .05$). Differences found in the number of OEs due to metering task (Baseline $M=.88$, $SD=1.13$; STA $M=.88$, $SD=1.13$ ) were non-significant ($F(1,12) = .000$, $p > .05$). No significant interaction was found between traffic density and metering task on number of operational errors.

Figure 12 could easily be interpreted as a simple correlation between aircraft count and operational errors, similar to effects in today's system. To shed further insight into this question, the operational errors during all short and medium runs were further analyzed on a sector-by-sector basis (Figure 13).

A one-way univariate ANOVA was performed and a significant effect was found for sector position (ZID80, ZID81, ZKC90, ZKC98) on the mean number of Operational Errors, ($F(3,204) = 9.00$, $p < .001$). Post-hoc pair-wise comparisons (Fisher's LSD) showed significant differences in mean number of OE to exist between the ZKC and ZID sectors (ZKC98 vs. ZID81, $p<.001$; ZKC90 vs. ZID81; $p<.001$; ZKC98 vs. ZID80, $p<.01$; ZKC90 vs. ZID80, $p<.01$).

No operational errors were observed in sector ZID81, which had an average of 30 aircraft, and peaked at 45 aircraft, while sector ZKC98 had the largest amount of



**Figure 11: Mean workload ratings over time, per sector, for the Short duration condition.**

operational errors, even though it typically contained much fewer aircraft than sectors ZKC90 and ZID81. Further investigation into the conflict locations revealed that the OEs seem to be related to the specific complexities mentioned at the beginning of this paragraph and occurred at sector boundaries and where uncoordinated dense departure streams interacted directly with arrival streams.
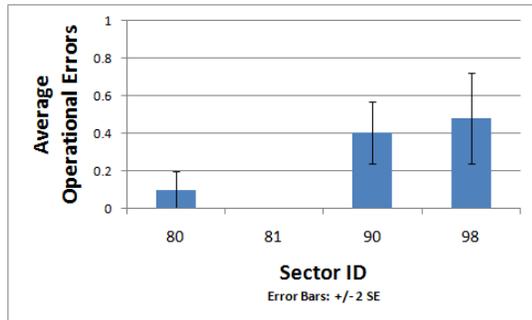


**Figure 23: Average Operational Errors by Sector**

### Qualitative results and discussion

Responses to post-run questions were analyzed using non-parametric repeated-measures statistics.

*Acceptability of separation assurance operations*

Among the post-run questions asked were six that 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 following questions were conditional upon previous answers. Participant answers can be compiled to form a scale from one to ten where one indicates that the operation is not safe and ten indicates that the system is acceptable. In Medium runs participants rated the separation assurance operation as safe less often under NextGen Traffic Level B (67.5% of the time) than under Level A (90.6% of the time) but although these differences were large they were not statistically significant. A general analysis of the comments indicated that it was not the absolute volume of traffic that concerned participants but that the situations became more complex as the traffic increased (and finding a conflict solution with a clear path became less and less easy). For example, one participant responded "Traffic level really not that much of a factor. Number of conflictions at any one moment were more of a factor."

In the short runs only the style of metering was varied under Level B traffic. Participants' safety ratings were very similar , here and differences were not seen in participants' ratings until the satisfactory question, where they said the separation assurance operation was satisfactory more often (97.2%) under the metering conditions with arrival time changes (S3 and S2) than under the static condition (S1) (83.7% of the time). This difference was not statistically significant.

There were significant differences in the short run CARS responses based on the sectors that controllers worked. Firstly, participants controlling ZKC sectors rated them as less safe ($\chi^2(3) = 7.909$, p=.048) and less satisfactory (($\chi^2(3) = 8.143,$ p=.043) more often than ZID sectors. When looking across the mean ratings for the four sectors, participants controlling ZKC98 rated it as safe only 61% of the time when they rated the other sectors as safe 85% of the time. and as satisfactory

only 71% when they rated other sectors as satisfactory 97.7% of the time. Overall, the ZKC98 sector separation assurance operations were rated as less acceptable than the operations in other sectors ($\chi^2(3) = 9.842,$, p=.02). From the general comments and other questions about workload, a reason for the lower ratings for sector ZKC98 is the greater number of events that occurred in this sector and that it had a more complex situation than the other sectors. This is in line with the qualitative findings from the medium runs.

In sum, the separation assurance operation was acceptable when there was less traffic but as the complexity of the situation increased (driven by greater volume) the operation approached a point at traffic level B when it became less than satisfactory or even unsafe for some sectors.

*Functional allocation between controller and automation*

In the post study questionnaire, one question listed sixteen different activities and asked which ones the participants would like to complete themselves and which tasks they would like the automation to take care of. The tasks ranged from "putting free track aircraft back onto their 4D routes" to "changing the range of the display" (see Figure 14).
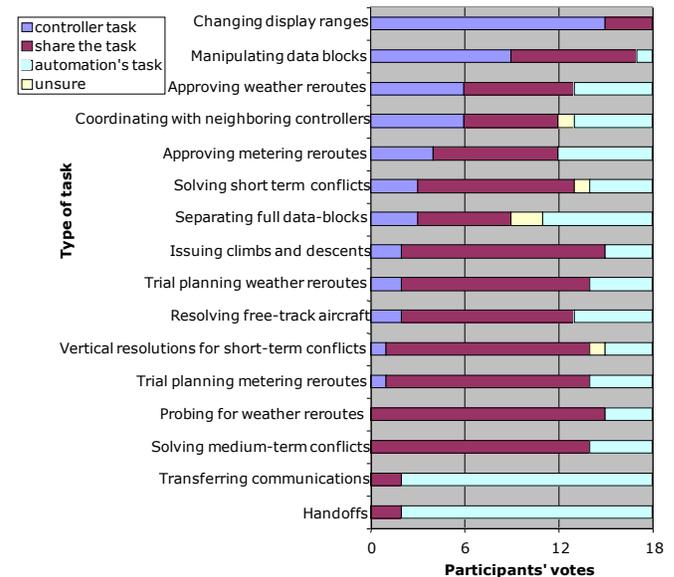


**Figure 34: Participants' allocation of tasks between the automation and themselves**

In general, participants most frequently voted that they should share the tasks with the automation. However, there were a couple of tasks for which they had a clear preference. Eighty-eight percent of participants (16) thought making and taking handoffs and transfer of communication should be left to the automation, and 83% of participants (15) thought the controller should have control of their display range adjustment. Twelve or more participants (66% plus) voted that six tasks or more should be shared between themselves and the automation. These tasks included trial planning, solving conflicts and issuing altitude changes for any reason. There were only three tasks that participants assigned fairly equally to themselves and the automation, these were "coordinating with neighboring sectors when conflicts span boundaries", "separating full data blocks" and "approving and sending reroutes due to weather". As general groupings, these

activities fall into routine tasks/ housekeeping and decision making tasks. Participants allocated the routine tasks to themselves or the automation but opted to share the decision making tasks.

A separate question asked participants to rate the level to which they had to trust the automation during the study. On average participants said they "had to trust the automation". However, perceptions of trust were intertwined with issues of responsibility. One participant for example responded "If the automation is responsible, then controllers should trust the system and not be held responsible. However the trust has to be there..." This is evidence that this is an important but complicated issue to address in NextGen environments.

*General participant opinions of the concept*

During debriefings participants were invited to comment about the concept and operations in general. Overall their comments were positive "you're on the right track" and "it seems fairly natural, why not do it?" Although participants were impressed at the level of simulation ("you're a lot further along than I thought"), they acknowledged that the technical realization of the concept still needs work: "…it seemed as if controller and automation fought against each other at times to resolve conflicts." Participants were diligent about cataloging problems but commented on the potential value in many of the tools, even that they would like some of these, such as the timeline functions, to be available today. One participant summed up the general feeling about the simulation with: "it's inevitable, I think the concept is strong, it needs work and testing, I think it's the way we're going to go."

## DISCUSSION AND CONCLUSIONS

The qualitative and quantitative results presented here lead to interesting preliminary conclusions. The overall traffic density has a primary impact on workload, safety and acceptability of the concept. The simulated NextGen traffic level A reflects the traffic density currently expected by the FAA for the year 2025. Level A was handled in all sectors with acceptable workload, very low operational error rates, and was largely found to be acceptable and safe by the operators.

Traffic Level B resulted for most sectors in unacceptable and unsafe operations. However, the fact that the highly loaded sector ZID81 (48 aircraft peak) was worked without operational errors while operations in the moderately loaded sector ZKC98 (32 aircraft peak) resulted in various OEs, points to a very promising effect. It appears that under the concept of ground-based automated separation assurance, safety, workload and acceptability are no longer directly linked to the total aircraft count within a given sector. This is very important, because it opens the door towards providing the capacity to accommodate much higher future airspace demand. The key seems to be to identify, eliminate or accommodate the particular complexities that made certain operations unmanageable.

Overall, the participants' comments indicated that this concept, with the right allocation of functions between the controller and the automation, has the potential to providing the airspace capacity required for NextGen.

## REFERENCES

1. Pearce, R.A. (2009) "The NextGen JPDO Model of Interagency Planning," AIAA 2009-7010, 9th AIAA ATIO, 21 - 23 September 2009, Hilton Head, SC.

2. Wing, D., Prevot, T., Murdoch, J., et al. (2010) "Comparison of Ground-Based and Airborne and Function Allocation Concepts for NextGen Using Human-In-The-Loop Simulations" AIAA, 10th ATIO, in preparation.

3. Concept of Operations for the Next Generation Air Transportation System Version 3.0. [Online] October 1, 2009. http://www.jpdo.gov/library/NextGen_ConOps_v3.pdf.

4. Erzberger, H. (2005) "Transforming the NAS: The Next Generation Air Traffic Control System," 24th ICAS, Yokohama, Japan, August 2005.

5. Wing, D. (2005) "A Potentially Useful Role for Airborne Separation in 4D-Trajectory ATM Operations," AIAA 2005-7336, 5th AIAA ATIO, 26-28 Sept 2005, Arlington, Virginia.

6. Prevot, T., Homola, J., and Mercer, J. (2008) "Human-in-the-Loop Evaluation of Ground-Based Automated Separation Assurance for NextGen," ICAS 2008-11.4.5, and AIAA-ATIO-2008-8885, Anchorage, AK.

7. Prevot, T., Homola, J., Mercer, J., Mainini, M. and Cabrall, C. (2009) "Initial Evaluation of NextGen Air/Ground Operations with Ground-Based Automated Separation Assurance," ATM2009, Napa, California.

8. FAA (2010) NAS Enterprise Architecture Portal https://nasea.faa.gov

9. Prevot T. 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.

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

11. FAA Aerospace Forecasts FY 2009-2025 http://www.faa.gov/data_research/aviation/aerospace_forecasts/20092025/media/FAA Aerospace Forecasts FY 2009-2025.pdf

12. Erzberger, H. and Heere, K. (2008) "Algorithm and operational concept for solving short range conflicts," ICAS 2008-8.7.5, Anchorage, Alaska, Sept 15-19.

13. Prevôt, T., Lee, P., Callantine, T., Mercer, J., Homola, J., Smith, N., and Palmer E. (2010) "Human-in-the-Loop Evaluation of NextGen Concepts in the Airspace Operations Laboratory," AIAA 2010-7609 MST, Toronto, Canada.

14. Lee, K., Kerns, K., Bone, R., and Nickelson, M. (2001) "Development and validation of the Controller Acceptance Rating Scale (CARS): Results of empirical research," ATM2001, Santa Fe, NM, December