Exploring Workload Factors Across Future Environments

Jeffrey Homola, Lynne Martin, Joey Mercer
NASA Ames Research Center/
San Jose State University Foundation
Moffett Field, CA 94035
+1 650 604 4063
jeffrey.r.homola@nasa.gov

Thomas Prevot
NASA Ames Research Center
Moffett Field, CA 94035
+1 650 604 2441
thomas.prevot@nasa.gov

ABSTRACT

A human-in-the-loop simulation was conducted that examined separation assurance across four progressive future time frames. Decision support, traffic density, separation assurance roles and responsibilities, and aircraft equipage mix were varied across conditions. In a near-term condition, these factors were set to approximate current day operations. In contrast, the most far-term condition involved two times current traffic, full air-ground data communications equipage, and automated conflict resolution working independently. The variation across the four conditions provided an opportunity to explore the pattern of reported controller workload, and what factors contributed to any observed differences. Despite increasing levels of traffic, results showed that mean workload ratings did not differ across conditions with the exception of the furthest term condition which was significantly lower. However, additional analyses were conducted that examined the relationship between workload and the varying traffic characteristics per condition. Although each condition had different significant contributors to workload, the one consistent contributor to workload across each condition was the number of conflicts. This result highlights the importance of work being done to develop the concepts and automation necessary to progressively balance the allocation of separation assurance functions between automation and the air traffic controller of the future.

Keywords

Human-in-the-loop, workload, separation assurance

INTRODUCTION

The National Airspace System (NAS) of today is a very complex system with any number of interacting and competing components at play at any given time. Within this system, the air traffic controller is responsible for ensuring safe separation of aircraft and providing service when able. While controllers are, for the most part, capable of performing these tasks today, it is becoming apparent that the NAS, as a whole, is approaching its limits with respect to available capacity to accommodate traffic demand. Compounding this matter is the continued projection by the Federal Aviation Administration (FAA) [1] of a significant increase in demand in terms of traffic as well as types of aircraft and varied levels of equipage expected to operate within the same airspace.

In order for the NAS to make the transition from today's environment to that which is envisioned for the future, it is evident that potential barriers or factors need to be identified and addressed. One such factor that has been identified is that of controller workload.

Within the air traffic control domain and beyond, there are differing opinions on what a formal definition of workload actually is. For the purposes of this paper, it is defined as the perceived mental and physical effort required for maintaining a safe and expeditious flow of traffic.

In that context, workload likely has a variety of influential dynamic and static factors that interact differently depending on the situation. Given that human workload capacity cannot simply be scaled to accommodate future scenarios, other means are necessary to expand the NAS beyond its current limits. However, in order to reduce or eliminate human workload as a limiting factor, a greater understanding of its contributors is needed.

The area of research into air traffic controller workload and the underlying airspace complexity is by no means new (references [2, 3] provide extensive reviews of the published literature). A common thread relating much of this research has been the role that simple aircraft count or traffic density plays in the development of workload. Although it is still used in operations today as a proxy for anticipated air traffic controller workload through the Monitor Alert Parameter (MAP), the value of attributing great weight to aircraft count has been called into question. In addressing this issue, researchers have included a number of additional variables –both existing and constructed– as a means to more fully capture the relationship between airspace and traffic factors with workload.

In this effort to examine a wider range of workload factors, one question that arises is that of how workload will be impacted as other factors scale toward future operational environments. To date, much of the work that has gone into the area of workload analysis and prediction has been accomplished through modeling, passive interviews, or part-task studies of current day operations. Through this paper, an attempt will be made to extend the analysis of workload and its relationship with other factors to a range of near to far-term future environments as simulated in a recent human-in-the-loop (HITL) simulation.

HITL DESCRIPTION

A HITL simulation was conducted in the Airspace Operations Laboratory at the NASA Ames Research Center that examined the impact of integrating self-separating aircraft into a ground-based separation assurance environment across progressively emergent stages of what is envisioned to be part of the Next Generation Air Transportation System (NextGen) [4, 5]. This design of simulating operations across successive stages of NextGen provided an opportunity to examine how the changes in each NextGen environment affected workload reported by the air traffic controller participants. A short description of the NextGen conditions simulated in the HITL follows in order to provide context for the ensuing workload results. For a more comprehensive description of the overall simulation, refer to [4, 5].

Design

A total of four NextGen conditions were simulated over the course of the study: Current Day, Minimum NextGen, Moderate NextGen, and Maximum NextGen. Each condition will be described in turn.

Current Day

The Current Day condition was the most near-term condition and approximated current operations with the addition of Automatic Dependent Surveillance-Broadcast (ADS-B) out surveillance data. All aircraft in this condition were not equipped with air-ground data communication (data comm) equipage and were referred to as unequipped. This meant that clearances and handoff instructions were required to be transmitted via voice as it is done today rather than uplinked messages. Aircraft were displayed using a yellow chevron as the target symbol and full data block while inside the sector. Controllers were required to perform manual handoffs and frequency changes on all aircraft entering and exiting their sectors with the appropriate exchange of verbal communications. Pilots of unequipped aircraft that were local arrivals were required to request a lower altitude upon reaching their top of descent. Likewise, departures of unequipped aircraft that would not reach their top of climb prior to entering the test area were pre-assigned a temporary altitude limit of flight level (FL) 320. It was then the controller's decision to allow those aircraft to climb into the sector to their filed or an amended altitude.

In this environment, traffic levels were designed to represent current day peak traffic levels with a Monitor Alert Parameter (MAP) value of 18 aircraft per sector.

Minimum NextGen

The Minimum NextGen condition introduced limited air-ground data comm equipage to the airspace. Approximately 25% of the simulated aircraft were made up of aircraft with data comm equipage that enabled automatic handoffs between sectors with a subsequent automatic transfer of communications. This capability meant that the controllers did not need to perform any actions or communicate with data comm equipped aircraft for performing handoffs, as they normally would. Equipped aircraft were displayed with grey chevrons as target symbols with limited data blocks by default. Equipped aircraft were cleared to descend at their top of descent point. To aid in anticipation of this event, the limited data block popped up to full when the aircraft was within 150 nautical miles of its destination airport. Additionally, the pilot of the equipped arrival aircraft notified the owning controller when leaving their current altitude. Equipped departure aircraft were also cleared for their climb without the need for controller approval. Unequipped arrival and departure aircraft were handled the same as in the previous condition.

This condition also introduced decision support tools (DSTs) for the controllers in the form of trajectory trial planning functionality as well as a conflict list. However, none of these tools were integrated with data comm, which meant that all control instructions still needed to be communicated via voice regardless of equipage.

The introduction of data comm and DSTs were thought to enable greater airspace capacity, and, subsequently, the MAP value for the test airspace was increased to 22 aircraft per sector as a first step away from current day and toward a NextGen environment.

Moderate NextGen

In the Moderate NextGen condition, expanded data comm capabilities were introduced that integrated with the controller planning tools on the ground and the flight management systems of the aircraft in the air. The ratio of data comm equipped and unequipped was increased such that approximately 50% of aircraft in the test airspace were data comm equipped. In addition to the tools available in the previous stage, automated conflict resolution support was introduced. Data comm in this stage enabled controllers to issue trajectory change instructions directly to equipped aircraft.

Traffic levels in this stage represented a 50% capacity increase over the Current Day levels, which translated to a MAP value of 27.

Maximum NextGen

The Maximum NextGen condition was the most far-term condition, and represented a significant departure from the previous three stages. The responsibility for conflict detection and resolution was assigned to the ground automation and all aircraft were data comm equipped. For conflicts detected, the ground automation computed trajectory-based resolutions and issued them directly to the aircraft, provided the computed resolutions did not exceed preset tolerances. If the resolution did exceed tolerances, the conflict was flagged to the controller for intervention. Given the technologies present in this stage, the role of the controller shifted from one of direct control to that of supervisory or management by exception. Such a shift was thought to enable a 100% increase in traffic levels over those in Current Day, which translated to a MAP value of 36 aircraft per sector.

Simulation Airspace

The simulated airspace consisted of five adjacent high altitude, en route test sectors (see Fig. 1). These sectors were assigned to two areas of specialization with sectors 26, 38, and 79 assigned to the North area and the remaining 49 and 59 to the South area. The floor of the overall test airspace was set at FL 330.

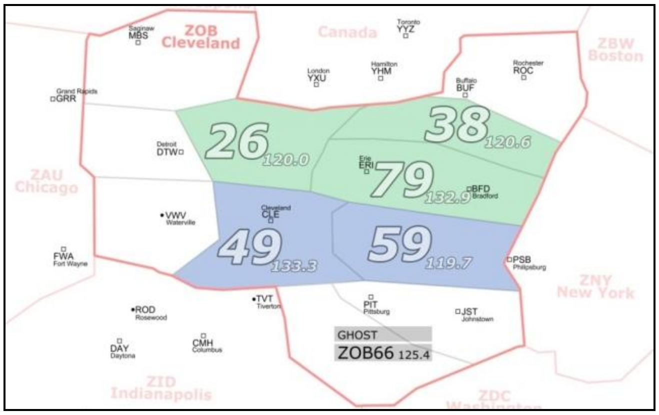

Figure 1. Test airspace within Cleveland ARTCC (areas of specialization denoted by color)

Traffic Characteristics

The traffic scenarios were based on actual traffic from the Cleveland Air Route Traffic Control Center (ARTCC) area, but modified to approximate the demand levels for each of the four NextGen stages as shown in Figure 2. Additionally, the equipage mixture of aircraft was designed to reflect each of the stages as outlined in the previous section (see Fig. 3). The overall traffic included a mix of level overflights as well as a number of arrivals and departures to and from area airports. Each scenario was designed for a 40-minute run length with the traffic building up gradually to peak around the midpoint of the run.

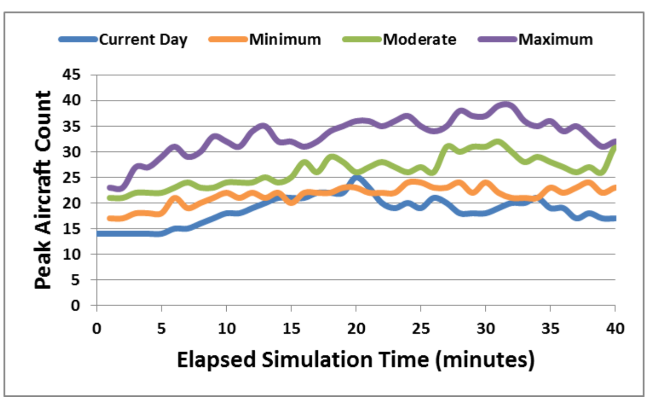

Figure 2. Peak traffic levels per condition

Through the design of the scenarios and the interactions of the controllers with the traffic, a number of conflicts occurred naturally. There was a varied mix of conflict types in terms of level and transitioning aircraft as well as the equipage mix of aircraft involved.

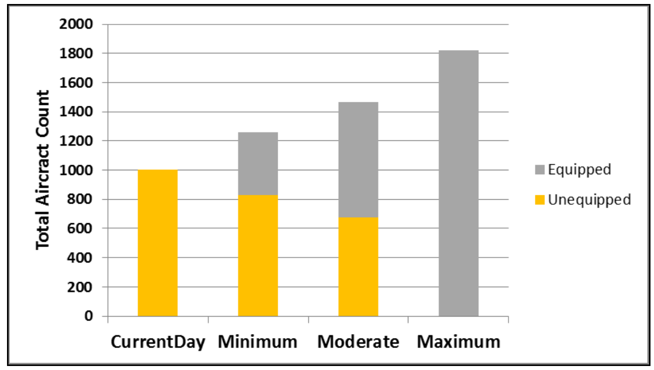

Figure 3. Traffic levels and equipage mix per condition

Apparatus

The simulation platform used in this study was the Multi Aircraft Control System (MACS) [6], which is being developed in the AOL, and has a wide range of simulation and rapid prototyping capabilities**.** Each controller workstation was equipped with a Barco display and Data System Replacement (DSR) trackball and keyboard. Voice communications were conducted through a custom, stand-alone voice system with a dedicated server.

Participants

A total of seven individuals served as test participants for this study. Six were current front line managers from various US ARTCCs, and one was a recently retired front line manager. Five of the test participants served as radar controllers and two as area supervisors. In support of the test participants, five retired controllers staffed radar associate positions. An additional three retired controllers acted as confederate “ghost” controllers responsible for traffic outside the test airspace. Ten airline pilots operated eight mid-fidelity, single-aircraft flight simulators, and ten general aviation/corporate pilots operated multi-aircraft stations.

The five test sectors comprised the North and South areas of specialization and were staffed in physically separate rooms. Each area had an assigned area supervisor that monitored the traffic situation as well as the workload of the participant radar controllers. It was the decision of the supervisor regarding when to provide radar associate support to the radar controller.

Procedure

Data collection occurred over the course of eight days. The four test conditions were presented to participants in successive order from Current Day to Maximum NextGen. A randomized or counter-balanced design was not pursued due to the confusion observed during simulation preparations when participants switched between different operating environments.

Two days were devoted to each condition, which consisted of training and six data collection runs. Each run was 40 minutes in length followed by an online, post-run questionnaire. Each sector controller participant remained at the same sector throughout the study for continuity and airspace familiarity.

DATA ANALYSIS APPROACH

The simulation of four successive NextGen stages provided an opportunity for the analysis of how workload might have changed in response to the changes in traffic levels, equipage, decision support, and procedures across the conditions. Of particular interest was which factors influenced changes in workload and in what way.

Workload Data

Throughout each run, an audible workload prompt was presented on the participants' display every three minutes. The response scale ranged from 1 ('very low') to 6 ('very high'). For part of the ensuing analysis, these ratings were further categorized according to Low (1, 2), Medium (3, 4), and High (5, 6).

Data Extraction and Treatment

Data were gathered from each of the 24 runs, from each of the five test sectors, and organized for extraction. The starting point for extraction was the reported workload data from each of the test sector participants. Each workload rating was extracted along with its time stamp and assigned the appropriate workload category. The time stamp of each workload rating served as the reference for subsequent extraction of the remainder of the data.

Since workload is not necessarily the reflection of an instantaneous traffic state, but rather the result of a build-up of various factors over time, the decision was made to relate selected factors to workload accordingly. Initially it was thought that a three-minute span for aggregating the other data over was the most practical approach since that matched the workload prompting schedule. However, after examining combined plots of various data over time with workload ratings overlaid, the decision was made to extend the duration of the time span from three minutes to five in order to potentially capture more of the underlying influences of factors on workload.

Based on this approach, the data for selected factors was aggregated into overlapping, five-minute bins relative to each workload rating. This meant that, for example, aircraft count data was derived for each associated workload rating by assigning the callsign of each unique aircraft owned by the controller to each relevant five-minute bin and summing each bin to get a resultant value for the number of aircraft owned.

It should be noted that the first workload rating of a run was discarded since data was unavailable for the entire five minute period. Additionally, it was felt that three minutes was simply too early into a run to consider airspace and operational data -as it relates to workload- useful.

RESULTS

Subjective Workload Results

Throughout the study, a total of 1,521 workload ratings were collected. An initial analysis was conducted in order to examine the mean differences in workload between each of the four experimental conditions. Figure 4 presents the results of mean workload where it can be seen that the first three conditions had nearly identical mean ratings (Current Day: *M*= 3.26, *SD*= 0.97; Minimum NextGen: *M*= 3.21, *SD*= 0.70; Moderate NextGen: *M*= 3.15, *SD*= 0.83), and that there was a sharp drop in reported workload in the far-term Maximum NextGen condition (*M*= 1.79, *SD*= 0.52).

A one-way, repeated measures Analysis of Variance (ANOVA) was conducted to determine whether there were any meaningful differences in the workload ratings where a significant main effect was found, $F(3, 12) = 32.43$, $p < .01$. Tukey's honestly significant difference (HSD) post hoc test showed that the mean workload in the Max NextGen condition significantly differed from the other conditions and that the others did not differ significantly from each other.

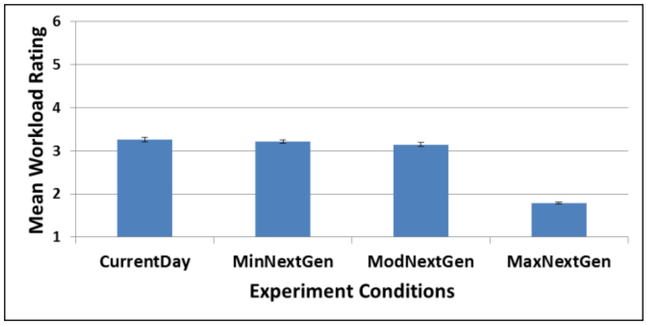

Figure 4. Mean workload ratings per condition

In addition to the mean workload ratings, the distribution of workload according to the categories of Low, Medium, and High in each of the conditions was examined. Figure 5 presents a breakdown of the percentage composition of each workload category per condition. It can be seen that the first three conditions were overwhelmingly reported as inducing Medium levels of workload whereas the Maximum NextGen condition was characterized by Low workload. The pattern of High workload appeared to follow a decreasing trend as the conditions progressed.

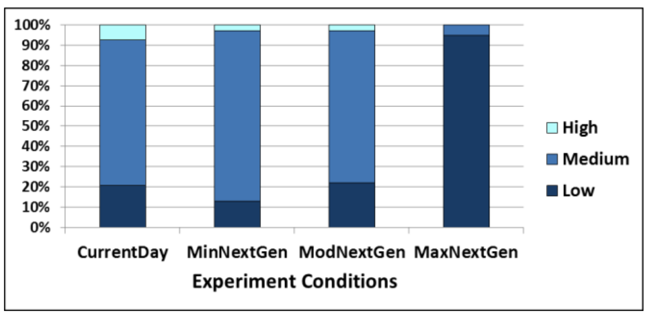

Figure 5. Percentage breakdown of workload by category per condition

Condition Differences

As described in the HITL Description section, the Maximum NextGen condition represented a significant departure from the others in terms of the controller's role shifting from that of direct control to one of automation supervisor and responsibility for separation assurance transitioning from controller to automation. The workload results thus far highlight this difference in operational paradigm in that workload in the Maximum NextGen condition was significantly lower than the other conditions and had a different composition of workload categories. As a result, the Maximum NextGen condition was excluded from further analysis.

Factors Affecting Workload

The results presented thus far speak to the generalities of workload in the context of different operational timeframes. However, they do not begin to provide any answers to the much studied question of what the drivers of workload are, and, more importantly, how those might change as roles and responsibilities change, and technologies are progressively introduced to the airspace environment.

To begin to address these questions, a number of factors were considered for inclusion in the overall workload model. A total of 18 factors were initially selected. Data for each of the selected factors was collected and calculated for distribution to the appropriate five-minute period relative to each workload rating time stamp.

Exploratory analyses of the data followed, with the intent of reducing the number of factors included for the final analyses. The purpose of this reduction was for relevancy, simplicity, and understandability of results. An additional purpose was to eliminate the amount of overlap between certain factors. Examination of relationships revealed strong correlations between a number of factors (Figure 6 presents a visualization of the correlation matrix), which led to the decision to exclude or combine those that were highly correlated with other similar factors as well as those that were sub-levels of larger categories (e.g., short-term conflicts and strategic conflicts were sub-levels of Total Conflicts).

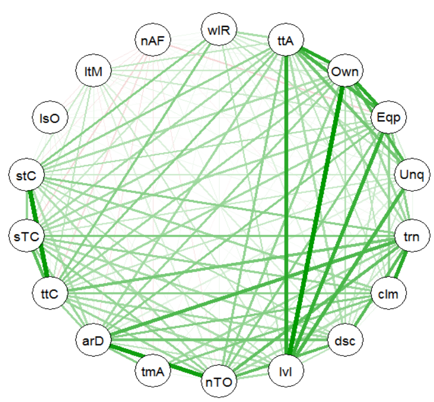

Figure 6. Visualization of correlation matrix (line thickness represents strength of relationship)

Based on this approach, the total number of factors was subsequently reduced from 18 to five: Equipped, Unequipped, Transitioning, Arrival-Departure, and Total Conflicts. An item to note here is that for the Current Day condition, the number of factors was reduced to four because the traffic was composed of unequipped aircraft.

The Equipped and Unequipped factors refer to the unique number of data comm equipped and unequipped aircraft, respectively, that were owned by the controller during the five minutes prior to a given workload prompt. The Transitioning factor represents the number of aircraft that were either in a climb or descent during that period. The number of aircraft that were either within 150 nautical miles of their destination airport or a local departure awaiting airspace access at a temporary assigned altitude composed the Arrival-Departure factor. Finally, the Total Conflicts factor was represented by the unique number of conflicts in which at least one aircraft in a conflict pair was owned by the sector controller or the predicted loss of separation (LOS) point was within a given sector.

Descriptive Analysis of Factors

Having selected the factors, an analysis of descriptive statistics was conducted that compared the values, within each factor, between the Low, Medium, and High workload ratings. Figures 7, 8, and 9 present the results of this analysis for the three conditions under examination.

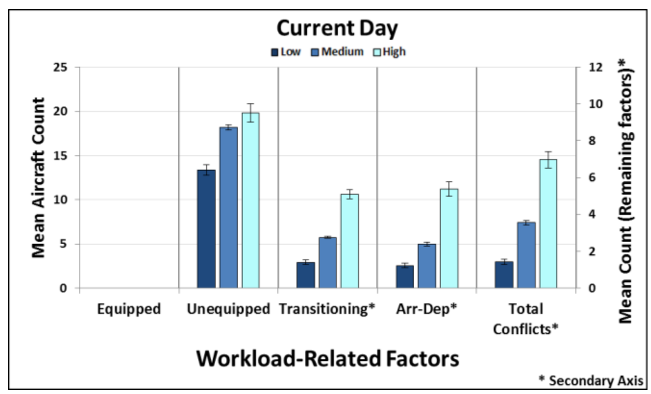

Figure 7. Comparison of means between workload categories for each factor in the Current Day condition

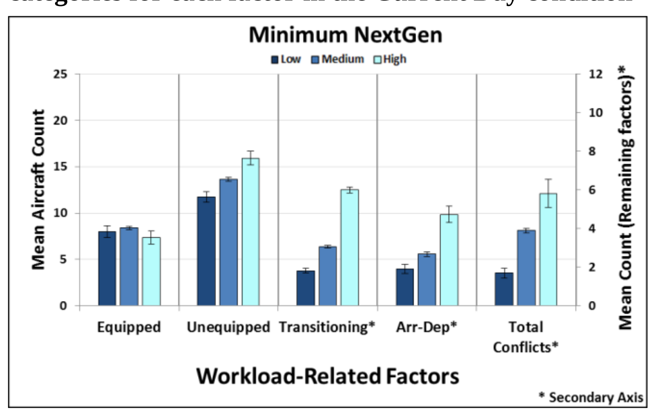

Figure 8. Comparison of means between workload categories in the Minimum NextGen condition

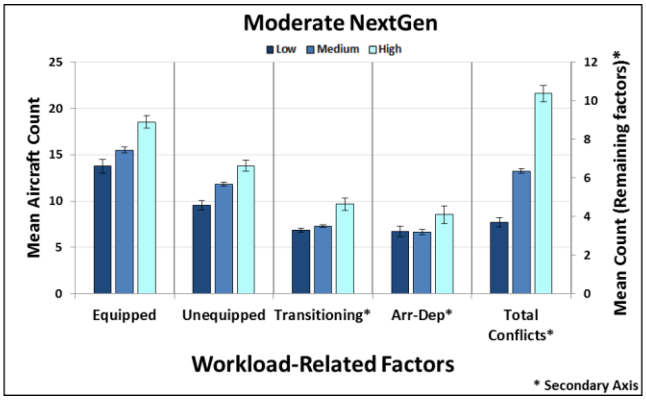

Figure 9. Comparison of means between workload categories in the Moderate NextGen condition

Results from the Current Day condition showed a fairly uniform and linear trend in the differences between Low, Medium, and High workload categorizations for each of the factors as one might expect (Fig. 7).

This trend continued, for the most part, in the results for the Minimum NextGen condition with the exception of the mean number of Equipped aircraft (Fig. 8). While the mean number of Unequipped aircraft was progressively higher across workload categories, the mean number of Equipped aircraft did not differ; in fact, there were fewer equipped aircraft in the test sectors in the lead up to high workload ratings relative to when ratings of Low and Medium workload were reported. This result speaks to the benefits of adding data comm equipped aircraft to the airspace in that throughput was increased without a subsequent increase in workload.

In the Moderate NextGen condition, however, the trend for equipped aircraft as observed in the Minimum NextGen condition did not continue (Fig. 9): the mean number of equipped aircraft increased linearly in line with the workload categories similar to results for Unequipped aircraft. Interestingly, the mean differences between workload categories for the factors of Transitioning and Arrival-Departure aircraft became more muted compared to the previous conditions where there were progressive increases.

Comparing the results across conditions, the two factors that appeared to consistently map to a linear increase in workload were the number of unequipped aircraft and the total number of conflicts. To build upon these results and further examine the relationships between each factor and reported workload, correlations were computed.

Correlations

To assess the relationship between each of the workload factors and reported workload ratings, Pearson's product-moment correlation coefficients were computed (Table 1).

In the Current Day condition, each of the factors was significantly correlated with workload where Total Conflicts showed the strongest positive linear relationship followed by the Transitioning factor. Interestingly, the number of Unequipped aircraft was shown to have the weakest relationship with workload out of the factor set.

	Current Day	Minimum NextGen	Moderate NextGen
Equipped		-0.09	0.13*
Unequipped	0.32*	0.17*	0.27*
Transitioning	0.46*	0.42*	0.14*
Arrival-Departure	0.44*	0.33*	0.08
Total Conflicts	0.54*	0.46*	0.46*
	n = 355	n = 357	n = 352

Table 1. Pearson's r correlations for each factor with workload (significance of *p*<.05 denoted by *)

In the Minimum NextGen condition, the strength of relationships between the Total Conflicts and Transitioning factors with reported workload became more pronounced relative to the other factors. As seen in the descriptive statistics for Minimum NextGen from Figure 8, the Equipped aircraft factor appeared to have a less defined pattern between workload categories compared with the other factors. This observation was further supported by the Equipped factor showing a non-significant relationship with workload. However, it is interesting to note that the correlation results show that there is actually a negative relationship between the number of equipped aircraft and workload in this particular condition. This means that, in general, as the number of equipped aircraft increased, there was an observed decrease in workload.

Correlations from the Moderate NextGen condition revealed an interesting change in relationships compared to the other conditions in which the strongest relationships with workload were found to be Unequipped aircraft and Total Conflicts. Transitioning aircraft was shown to have a relatively weak relationship with workload in contrast to very strong relationships in correlations in the other conditions. Similarly, the Arrival-Departure factor was not significantly correlated with workload.

Multiple Linear Regression Analysis

To further examine the relationship between the selected factors and workload, a multiple linear regression was used to develop a model to predict workload from the set of predictor variables examined thus far: Equipped, Unequipped, Transitioning, Arrival-Departure, and Total Conflicts. Table 2 presents a summary of significant predictors that resulted from the regression models.

Current Day

In the Current Day condition, the resulting model provided the following function:

$$\hat{y} = 2.27 + 0.02x_{\text{uneq}} + 0.03x_{\text{trans}} + 0.10x_{\text{arr-dep}} + 0.13x_{\text{totalConf}}$$

Results from the analysis showed that 36% of workload variance could be explained by the four predictors (R^2= 0.36, F(4, 350)= 49.77, p< .001), and that Arrival-Departure (p< .001) and Total Conflicts (p< .001) were significant predictors. To address concerns of multi-collinearity between predictors, variance inflation factors (VIF) were computed. All VIF values were low (< 2.2) suggesting a low degree of multi-collinearity.

Minimum NextGen

In the Minimum NextGen condition, the Equipped factor was added to the regression model, which produced the following function:

$\hat{y}$= 2.92-0.04x_{eq}+0.01x_{uneq}+0.05x_{trans}+0.03$x_{arr\text{-}dep}$+ .09$x_{totalConf}$

Results from analysis of the five predictor model showed that it could explain 29% of workload variance (R^2= 0.29, $F(5, 351)$= 28.63, p< .001), and that the Equipped (p< .001) and Total Conflicts (p< .001) predictor variables were strongly significant while the Transitioning predictor variable was significant but to a lesser extent (p< .05). VIFs were also computed and were all shown to be relatively low (<2.4), again suggesting a low degree of multi-collinearity.

Moderate NextGen

A multiple linear regression was also performed for the Moderate NextGen condition where the model produced the following function:

$\hat{y}$=2.23+0.01x_{eq}+0.04x_{uneq}-0.03x_{trans}-0.03$x_{arr\text{-}dep}$+ .11$x_{totalConf}$

Based on this five predictor model, it was found that 26% of the variance in workload could be explained (R^2= 0.26, $F(5, 346)$= 23.93, p< .001), and that the Unequipped (p< .001) and Total Conflicts (p< .001) were strongly significant predictor variables. VIFs again showed a low degree of multi-collinearity between predictors in which all values were less than 2.3.

Current Day	Minimum NextGen	Moderate NextGen
Total Conflicts	Total Conflicts	Total Conflicts
Arrival-Departure	Equipped	Unequipped
	Transitioning	

Table 2. Significant predictor variables found through multiple linear regression per condition

DISCUSSION

A human-in-the-loop simulation was conducted that investigated the allocation of separation assurance functions across four progressive time frames. The design of this simulation, with its approximation of different emergent phases of NextGen, allowed for an exploration of how controller workload scaled in response to changes in airspace factors.

Overall Workload

A comparison of mean reported workload across conditions did not yield significant results with the exception of the furthest term condition, which was excluded from this analysis due to its significant departure in operations and procedures compared with the other three. Considering workload from the first three time frames, the fact that workload did not appear to increase in response to increases in virtually all airspace related factors not only speaks to the benefits that envisioned support tools and procedures may provide, but also supports the notion that increases in traffic count alone are not an inherently useful indicator of workload. This is particularly true moving forward into NextGen where the potential contributions of traffic count to workload are offset by advances in automation support and function allocation.

The categorization of workload into Low, Medium, and High provided the first glimpse of workload patterns across the different conditions. While much of the reported workload fell into the Medium category in each condition, High workload ratings differed. The greatest number of High workload ratings was observed in the Current Day condition with a noticeable drop in the Minimum and Moderate NextGen conditions. This result, again, highlights the benefits of introducing controller support as a means of removing potential barriers to the demands of a future system.

Selected Factors and Workload Relationships

Although mean workload did not show significant differences across conditions, the results masked the differences in impact that various factors may have had on whether ratings fell within the Low, Medium, or High categories. In addressing this aspect of the analysis, a particularly simplified set of workload factors was selected for examination: Equipped, Unequipped, Transitioning, Arrival-Departure, and Total Conflicts. This simplification was intentional in order to limit the interpretation of results to simple and actionable factors.

Current Day

In the Current Day condition, the mean workload across each of the categories was quite linear for each of the factors. An analysis of correlations supported this result in that each factor showed a significant positive correlation with workload with Total Conflicts showing the strongest relationship. Taking these factors into consideration as potential predictors of workload, a multiple regression analysis was conducted, which resulted in the Arrival-Departure and Total Conflicts factors being identified as significant predictors of workload. In this case, the number of aircraft either waiting for airspace access or nearing their top of descent was predictive of workload as was the number of conflicts presented to the controller.

Minimum NextGen

In contrast to the Current Day condition, the Minimum NextGen condition included the introduction of data comm equipped aircraft as part of a 25% increase in traffic count as well as trial planning capabilities. Figure 8 presents the mean values of each factor according to workload category where it can be seen that the values for Equipped did not increase across categories while the other factors did. In fact, correlation results showed that the Equipped factor did not have a significant correlation with workload. However, in the presence of the other factors as tested through the multiple linear regression, the Equipped factor proved to be a significant predictor. The negative coefficient in the model's function suggests that higher numbers of equipped aircraft in the sector tended to result in lower workload. This result is not surprising given the fewer actions required for control of equipped aircraft. However, it does speak to the need for considering the contributions of aircraft according to equipage differently. The Transitioning and Total Conflicts factors were also found to be significant predictors suggesting that greater numbers of climbing and descending aircraft as well as aircraft in conflict increased workload. The Unequipped and Arrival-Departure factors, on the other hand, did not appear to be significant predictors of workload in this condition.

Moderate NextGen

The Moderate NextGen condition involved a 50% increase in traffic with an equipage mix within the test sectors of 50% equipped and 50% unequipped. Conflict resolution support was also added. In terms of workload, unlike the Minimum NextGen condition, the greater proportion of equipped aircraft in the test airspace appeared to have an additive impact on workload, similar to the effects of unequipped aircraft (Fig. 9). An interesting difference in this condition compared with the others was that despite an increase in the overall traffic density in the airspace, the relationship between transitioning aircraft and arrival-departure aircraft with workload appeared to weaken such that the mean differences between workload categories was negligible and the correlation values were much lower. Results from the multiple regression further highlighted this result in that they were not significant predictors of workload, whereas the Unequipped and Total Conflicts factors were. Interestingly, it was in this condition that traffic count, at least with respect to unequipped aircraft, showed a stronger positive relationship with workload relative to the other factors (with the exception of Total Conflicts).

Summary of Findings

Having examined the relationship between various factors and workload across three different time frames, the strongest and most consistent predictor of workload was Total Conflicts. This consistency highlights the importance of work that has gone into the development of automation, support tools, and concepts for separation assurance to resolve conflicts as well as advancing support for traffic management to aid in the strategic avoidance of conflicts from occurring in the first place.

The other factors were not as straight forward or consistent as Total Conflicts in their ability to predict workload. In the Current Day and Minimum NextGen conditions, the numbers of aircraft in transition or nearing a transition as an arrival or departure were found to be significant predictors of workload while the number of equipped aircraft actually predicted lower workload. However, the number of unequipped aircraft served as a better workload predictor in the Moderate NextGen condition.

Finally, with respect to traffic count and its value as a workload predictor, it did not appear to significantly relate to nor predict workload until the Moderate NextGen condition. In this case, the Unequipped factor was the significant predictor whereas it was not for Current Day or Minimum NextGen. The reason for this result is unclear, but perhaps the combination of a higher traffic count and mix of traffic interacted to drive workload more so than in the other conditions.

Next Steps

The approach taken to examine workload-related factors in this analysis was fairly straightforward and meant to serve as a first step. Next steps to consider are the application of non-linear regression techniques to account for a larger portion of workload variance or approaching the workload data as a classification problem. Additionally, cross-validation techniques can be applied to the regression models as a means of assessing their predictive accuracy.

ACKNOWLEDGMENTS

The overall experiment from which this paper was derived was sponsored by the NASA Airspace System Program, Concept and Technology Development Project.

REFERENCES

1. Federal Aviation Administration. *FAA Aerospace Forecasts Fiscal Years 2014-2034*. Retrieved from http://www.faa.gov/about/office_org/headquarters_offices/apl/aviation_forecasts/aerospace_forecasts/2014-2034/media/2014_FAA_Aerospace_Forecast.pdf.
2. Hilburn, B. *Cognitive Complexity in Air Traffic Control: A Literature Review*. Brussels, Belgium: EUROCONTROL, 2004.
3. Loft, S., Sanderson, P., Neal, A., and Mooij, M. *Modeling and Predicting Mental Workload in En Route Air Traffic Control: Critical Review and Broader Implications*. The Journal of the Human Factors and Ergonomics Society June 2007 vol. 49 no. 3, 376-399.
4. Wing, D., Prevot, T., Lewis, T., Martin, L., Johnson, S., Cabrall, C. et al. *Pilot and Controller Evaluations of Separation Function Allocation in Air Traffic Management*. Proceedings of the 10th USA/Europe Air Traffic Management Research and Development Seminar, Chicago, June 2013.
5. Mercer, J., Prevot, T., Homola, J., Cabrall, C., Martin, L., Morey, S., and Gomez, A. *Human-Automation Cooperation for Separation Assurance in Future NextGen Environments*. Submitted to HCI-Aero 2014.
6. Prevot, T. *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.