RELATIONSHIP OF MAXIMUM MANAGEABLE AIR TRAFFIC CONTROL COMPLEXITY AND SECTOR CAPACITY

Parimal Kopardekar*, Jessica Rhodes, Albert Schwartz**,
Sherri Magyarits**, and Ben Willems****
*NASA Ames Research Center, Moffett Field, CA 94035,
****Federal Aviation Administration William J. Hughes Technical Center, Atlantic City, NJ
08405**

Abstract

In today's U.S. airspace operations, each sector has a capacity threshold in the form of the maximum number of aircraft. This threshold, called as Monitor Alert Parameter, serves as a controller workload limit indicator for each sector. Studies have shown that a number of factors, beyond a simple aircraft count, affect sector complexity and controller workload. This study examined if a sector's capacity could be based on a measure of maximum manageable complexity rather than a simple aircraft count. The results of a human-in-the-loop simulation study were used to identify the maximum acceptable complexity and its relationship with sector capacity. Aircraft in the study continued to enter the simulated sector until the controller participant terminated the simulation due to excessive workload, thus providing a sector complexity threshold. The number of aircraft at that complexity threshold varied, which provided a range of capacity values. In most cases, the controllers were able to handle more than the current instantaneous capacity limit. However, at higher complexities, losses of separations were observed. At the safest complexity range (1 to just under 4), where there were no separation losses, the maximum average number of aircraft controlled was about 24 indicating that sector capacities will vary if they are based on complexity.

1 Introduction

Current U.S. airspace operations use the Monitor Alert Parameter (MAP) as an indicator of a capacity limit for each sector. A sector's MAP value reflects maximum instantaneous aircraft count that can be safely handled by a sector controller. Currently, the MAP value is a function of the average time it takes for an aircraft to transition the sector [1]. In other words, MAP values are set to reflect controllers' acceptable workload. It is a straightforward and operationally intuitive method to keep the controller workload at a manageable level. It has long been understood that ten aircraft in a single line impose different complexity and workload than ten aircraft that are converging towards each other.

Prior research has shown that a number of factors affect complexity and air traffic controller workload. These factors include, but are not limited to, the number of aircraft and potential conflicts, number of hand-offs, heading and speed variation between two or more aircraft, aircraft proximity to each other, and presence of weather [2-5]. Controller workload is a subjective attribute and is an effect of air traffic complexity, which can be measured objectively. Both U.S. and European aviation researchers have been interested in developing quantifiable metrics for air traffic complexity that reflect controller workload. Studies have shown that complexity models perform better in representing controller workload than the MAP [4-6]. Yet, to date there has been no formal attempt made to establish sector capacities based on maximum manageable complexity.

Therefore, it would be valuable to study the airspace capacity limits based on maximum manageable complexity instead of MAP values. The overall objective of the complexity analysis reported in this paper was to examine if complexity could be used to set a sector's capacity limit. The specific goals of the analysis were threefold:

- Demonstrate if a complexity model better reflects the controller reported workload than an aircraft count-based model, like MAP.
- If the complexity model better reflects the controller workload, then determine sector capacity limit based on the complexity rather than a fixed aircraft count.
- Examine if the sector's complexity based capacity is different than the MAP value based capacity.

Although the first goal has been addressed in Kopardekar et al [6], in order to address the second and the third goals, it is important to verify that the complexity based model is indeed better than the aircraft count-based model.

2   Approach

The complexity analysis was performed using data from a human-in-the-loop (HITL) simulation of a Future En Route Workstation (FEWS) performed at the FAA William J. Hughes Technical Center in late 2006 – early 2007.

The following sections provide descriptions of the simulated environment, test conditions, and data collected, followed by the results of the complexity analysis of the relationship of maximum manageable air traffic control complexity and sector capacity.

2.1   Participants

Twelve controllers from Level 11 and 12 Air Route Traffic Control Centers (ARTCCs) participated in the HITL simulation. All participants were non-supervisory certified professional controllers qualified at their facility with a current medical certificate.

2.2   Scenarios/Airspace

The simulation consisted of 60-minute traffic samples, where the number of aircraft under control continuously increased. Increasing the traffic load steadily allowed for a better assessment of the relationship between task load (e.g., number of aircraft) and controller workload. Each scenario started with the controller being responsible for five aircraft and increased to as high as 50 aircraft at 50 minutes into the scenario.

The HITL study depicted a generic en route sector ZGN08 (Fig. 1). The MAP value of sector ZGN08 was assumed to be 21 based on an average aircraft transition time of 12 minutes (resulting in a MAP value of 18 ± 3) [1]. The complexity analysis focused only on Sector ZGN08 because the purpose of this study was to identify the maximum manageable complexity. Sector ZGN08 controllers were instructed to stop the scenario when they felt that maximum manageable workload, a subjective construct that represents complexity, was reached.

The airspace was designed by researchers and Subject Matter Experts (SMEs) to be realistic; containing several sectors, fix posting areas, Terminal Radar Approach Control (TRACON) facilities, navigational aids, airways, Standard Terminal Arrival Routes (STARs), and Standard Instrument Departures (SIDs), but also relatively easy for controllers to learn [7].
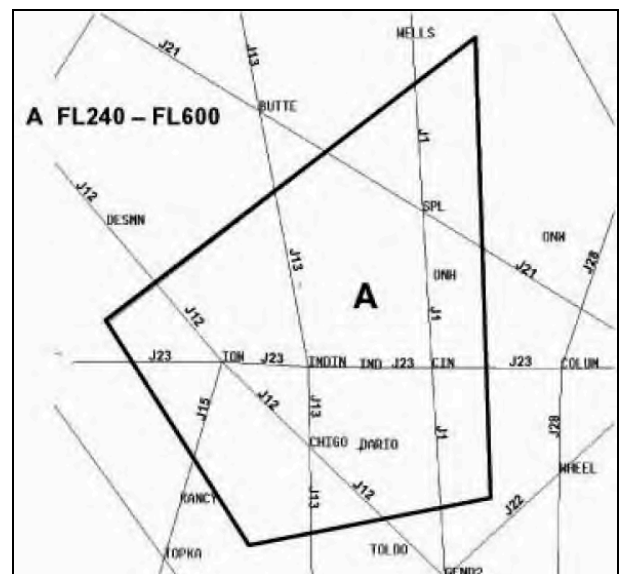

Fig. 1. Sector ZGN08 airspace.

2.3 Simulation Environment

In the one-sector configuration (two-person sector), controllers had a 2048x2048 pixel Barco Liquid Crystal Display available on the R-side that depicted the traffic situation (Fig. 2). Touch panels displayed electronic flight progress strips (eFPSs). Two touch displays showed incoming and current aircraft. Fig. 2 shows two passive eFPS panels on the left hand side. Next to the panels, the controller had a radar display, keyboard, trackball, and Keypad Selection Device (KSD) available. To the right of the Display System Replacement (DSR) radar display was the 2048x2048 display with a 1024x1024 inset, keyboard, trackball, and KSD for the D-side.

The simulation provided DSR functionality on the R- and D-sides. Most of the frequently used fielded functions were available except for the annotation (drawing) function. The D-side controller had a D-side Computer Readout Display (CRD) available for data entry and feedback information.

Fig. 2. Controller workstation

The hardware configuration used for the En Route Automation Modernization (ERAM) two-person sector emulation was identical to that used for the DSR emulation. The main differences between the ERAM system and the DSR/Host system were the hardware and software that processed incoming data and enabled controllers to interact with the system. The changes to the Computer-Human Interface (CHI) were mostly part of an update scheduled for DSR, which are now part of the ERAM deployment. These changes may affect the way controllers work with the system.

The hardware configuration used for the FEWS configuration had mostly the same hardware available as under the ERAM conditions except for the pointing device they used. In the FEWS conditions, the controllers had a 3-button Logitech Wheel Mouse instead of a trackball. The D-side controllers could still execute functions to assist the R-side controller but also had the full 2048x2048 radar display available. Controller interactions with FEWS presented the biggest differences from both ERAM and DSR conditions. Controllers had the ability to perform most functions using drop-down menus on the displays, as opposed to having to use the keyboards. In addition, FEWS provided automated datablock arranging, interactive electronic flight strips, ability to highlight aircraft on other displays, and automatic changes to D-side displays based on R-side actions.

2.3.1 Communications Configuration

The simulated communications system had links between the controller, Subject Matter Expert (SME) observer, simulation pilots, and experimenters and Push-to-Talk recording capability.

2.3.2 Simulation Pilot Terminal Configuration

A network linked six simulation pilot operator displays with the controller workstation positions.

2.4 Experimental Details

The experiment included three interface conditions; Display System Replacement (DSR), ERAM, and Future En Route Work Stations (FEWS-II); two communication configurations (CCs) Voice Communication (VC) or Data Communication (DC); and two staffing configurations (SC) options (2-person sector or 1-person sector). Both DSR and ERAM feature the same hardware configurations and displays, and both workstations give the D-side a limited set of capabilities to assist the R-side. However, DSR and ERAM differ in the processing of incoming data and some of the Computer-Human

Interface features (flight plan data and multiple flight strip readout).

In contrast, the FEWS workstation allows the D-side to have a more complete set of capabilities, which allow them to more directly, assist the R-side. While the hardware configuration of FEWS is mostly the same as that of ERAM, its Computer-Human Interface included automated datablock arranging, interactive electronic flight strips, the ability to highlight aircraft on other displays, and automatic changes to D-side displays based on R-side actions.

Controllers used the Traffic Management Advisor (TMA) decision support tool in all conditions. They were presented with a version of the National Airspace System (NAS) that had a mixed fleet of 70-75% aircraft with state-of-the-art equipment expected for 2015, whereas the other 25-30% still used equipment from 2005. Fully equipped aircraft had maximum automated flight capabilities including data communication, whereas the other aircraft continued to use 2005 capabilities.

2.5 Workload Rating

Workload ratings (on a scale of 1-10, with 1 being very low, and 10 being very high) were collected every two minutes using the Workload Assessment Keyboard (WAK) devices.

2.5.1 *Workload Assessment Keypad*

The WAK is a reliable and unobtrusive real-time, on-line measure of subjective workload. For the study, workload was defined as all the physical and mental effort one must exert to do his/her job. This included maintaining the "picture," planning, coordinating, decision-making, communicating, and whatever else was required to maintain a safe and expeditious traffic flow.

Each participant provided a workload rating on the WAK device every two minutes throughout a scenario. The system prompted participants to respond by emitting several beeps and illuminating the keypad buttons. The participants pressed one of the keypad buttons labeled from 1 (extremely low workload) to 10 (extremely high workload). At the low end of the scale (1 or 2), workload was low – the

controller could accomplish everything easily. Numbers 3, 4, and 5 represented increasing levels of moderate workload where the chance of error was still low but steadily increasing. Numbers 6, 7, and 8 reflected relatively high workload where there was some chance of making errors. At the high end of the scale were numbers 9 and 10, which represented very high workload, where it was likely that the controller would have to leave some tasks unfinished.

The participants had 30 seconds to respond; otherwise, the WAK recorded a code indicating that it received no response. Also, controllers were instructed that their ratings should reflect how much workload they were experiencing during the instant they were prompted to make the ratings.

2.6 Complexity Variables

In addition to staff configuration, communication configuration, and workstation types; 52 complexity variables [6] were also considered as independent variables for linear regression analysis.

3 Analysis

The following analysis approach was used:

1. Identify the variables that contributed to the complexity (see section 4.1).
2. Assess whether a model based on complexity variables represented workload better than a model based on an aircraft count (see section 4.2).
3. Record the numbers of aircraft that controllers were managing at the time the scenario was terminated (see section 4.3).
4. Compute complexity values (i.e., maximum manageable complexity) using the complexity-based model, when controllers terminated the scenario (see section 4.4).
5. Determine the sector capacity based on maximum manageable complexity when the scenarios were terminated (see section 4.5).
6. Determine maximum manageable capacity without any losses of separation (see section 4.6).

4   Results

4.1   Identification of Complexity Variables

The 52 complexity variables identified from prior complexity analyses were considered in this current analysis. A detailed list and description is provided in [5] and [6]. A step-wise linear regression was conducted with workload rating as a dependent variable and complexity variables as independent variable to identify those complexity variables that were significant. Table 1 lists the resulting significant complexity variables. Additionally, variables with a Variance Inflation Factor (VIF) value of 5.0 or higher were omitted due to inter-correlations.

4.2   Comparison of Aircraft Count-based Model and Complexity-based Model

For the complexity-based model, shown in Table 2, the coefficient of determination (R^2) was 0.74 (R = 0.86).

A regression was also conducted with workload rating as a dependent variable and aircraft count, staffing and communication, as independent variables. The coefficient of determination (R^2) was 0.65 (R = 0.81). It shows that the complexity-based model better represents workload than the aircraft count based model. Both the complexity-based model and aircraft count-based model included staffing and communication configuration as additional variables for consistency.

Fig. 3 shows the relationship of complexity, as computed from the regression model, and controller workload ratings for a range of aircraft. It shows how the predicted complexity values track the workload ratings.

Table 1. Statistically Significant Complexity Variables

| Variable	Estimate
Intercept	1.19
Staffing	0.24
Communication	0.48
Aircraft density by sector	2265.26
Inverse weighted horizontal separation	-47.51
Inverse minimum vertical separation	0.02
Variance of groundspeed	0.02
Altitude Variation	0.00
Number of aircraft close to sector boundary	0.11
Convergence angle between aircraft in conflict	-0.30
Number of altitude changes	-0.05
Number of aircraft within 10 miles of sector boundary	-0.95
Aircraft heading difference with respect to the heading of major axes of sector	0.03
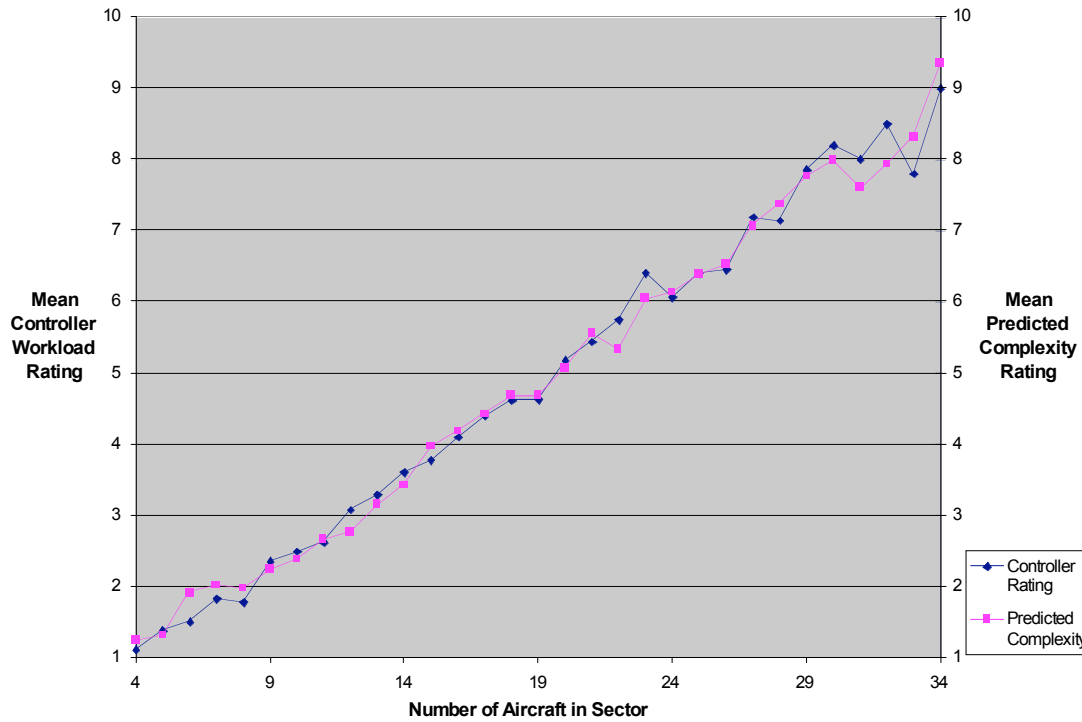
Fig. 3. Complexity values and controller ratings across a range of aircraft count.

4.3 Maximum Number of Aircraft Managed

Fig. 4 is a histogram of controller workload ratings for the last entry, just prior to scenario termination. It indicates that controllers' last workload ratings were mostly between 5 and 10, indicating that workload was usually between moderate and high when the scenario was terminated. This reinforces the notion that most controllers decided to terminate the run at higher workload values. Clearly, there are interpersonal differences, as some controllers decided to terminate at moderate workload and others decided to wait until very high workload. There are also four instances where controllers terminated the run after a low workload rating (e.g., 2, 3, or 4), which could

indicate an extremely proactive approach or a conservative interpretation of the rating scale.

4.4 Maximum Manageable Complexity

Fig 5 represents the predicted complexity value for the last workload entry. By comparison, the minimum complexity value at the last entry was higher than the minimum workload rating, and more complexity values than workload ratings are in the moderate to high range. In addition, there are some low complexity values at the last entry, which may indicate a proactive approach where controllers stopped the run before the complexity of the traffic became overwhelming.
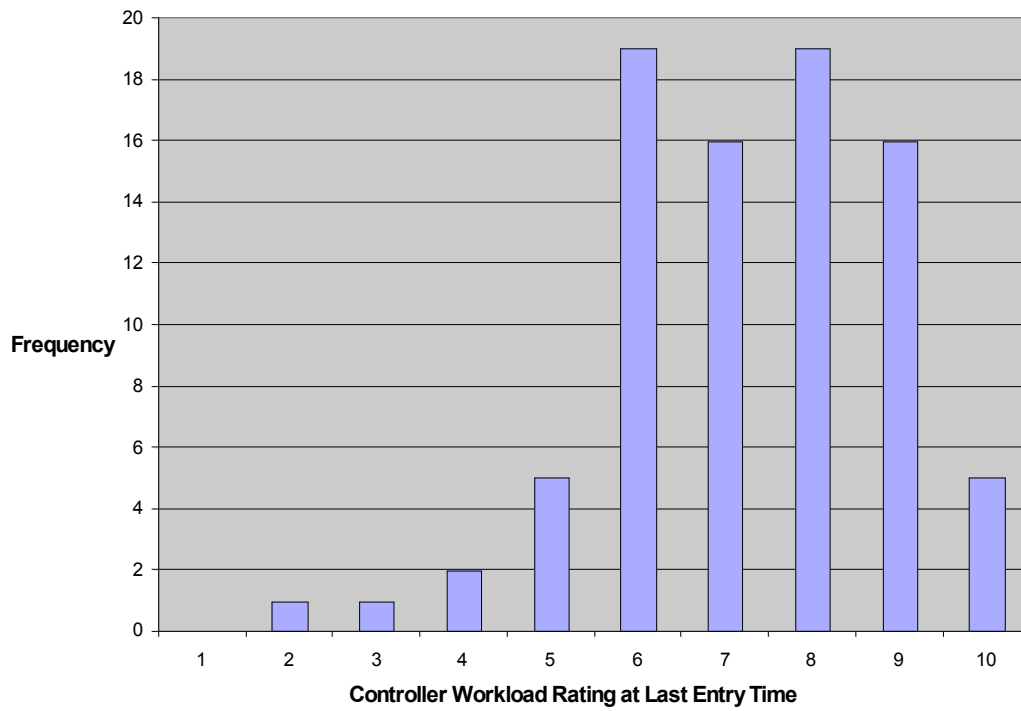
Fig. 4. Controller workload ratings just prior to scenario termination.
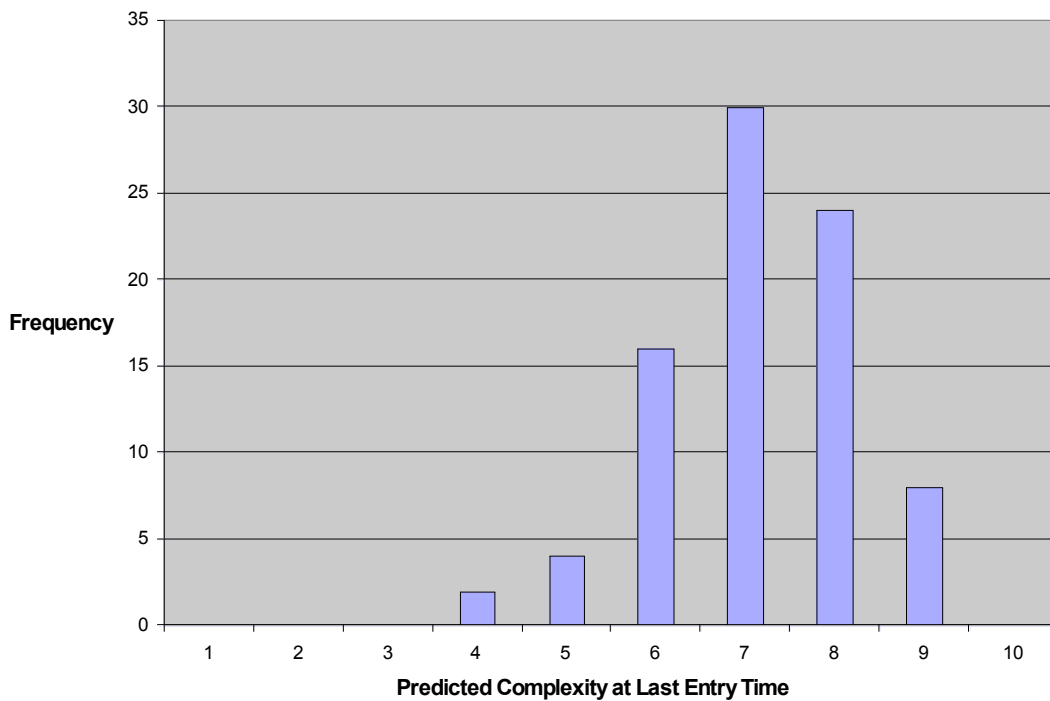
Fig 5. Last complexity values

Maximum manageable complexity was calculated based on when the controllers decided to terminate the scenario. This will correspond to their maximum manageable workload ratings. The complexity rating just prior to their stopping the scenario was considered as the maximum manageable complexity. Fig. 6 indicates that the average complexity when the controllers decided to stop the scenario ranged from just over 5 to just below 9. Although most of the time controllers stopped the scenario when there were more than 21 aircraft, there were a few instances
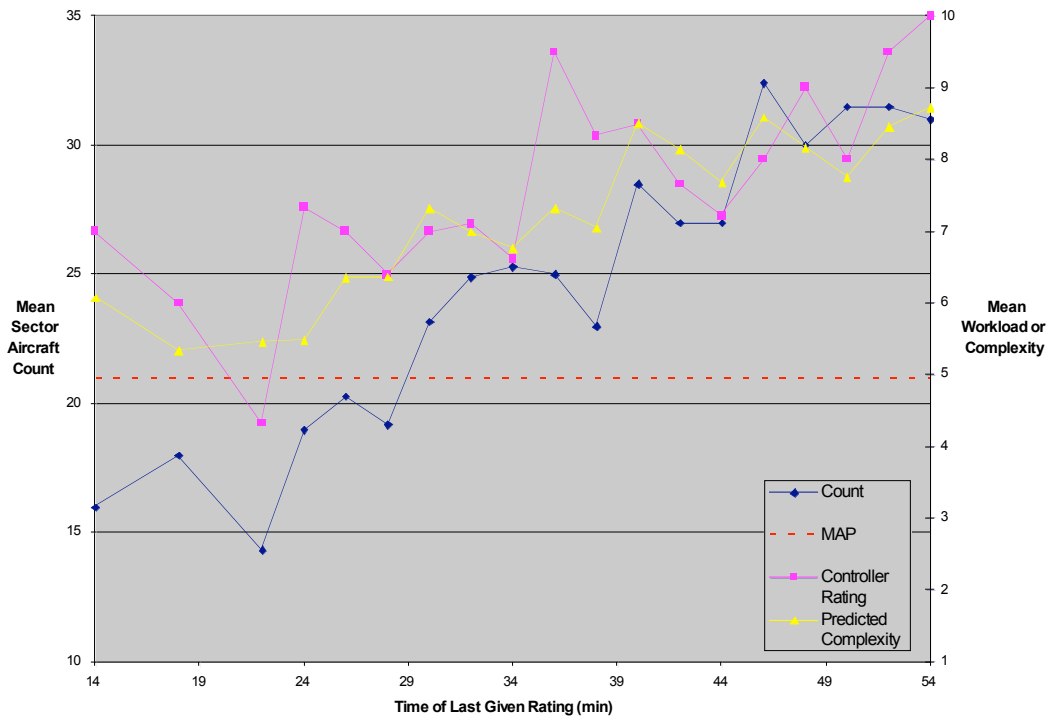
Fig. 6. Average aircraft count, average workload rating, and average predicted complexity value at last entry.

when they stopped the scenario with an aircraft count below 21. This implies that if complexity were to be used for setting up capacity limits, it is likely varying capacities would result, but mostly higher than the MAP. This also supports the claim that complexity, and not aircraft count alone, should be used as a workload and capacity metric. In some instances the controllers terminated the scenario earlier with lower complexity and a lower number of aircraft than the MAP value. This may be because of the differences in their styles; they were being proactive and were anticipating very high workload later, so they decided to terminate early.

4.5 Assessment of Airspace Capacity based on Maximum Manageable Complexity

It is interesting to note that the complexity threshold corresponding to the MAP value was about 5. It appears that by using complexity as an indicator, higher sector capacity limits could be achieved in most cases. The maximum complexity range was between 5 and 9. The challenge was to identify a safe range of complexity values, which would increase the capacity but not significantly decrease the safety. Figures 7 and 8 show that the perceived predicted complexity was lower for data link scenarios, under different display configurations, as compared with the voice scenarios.
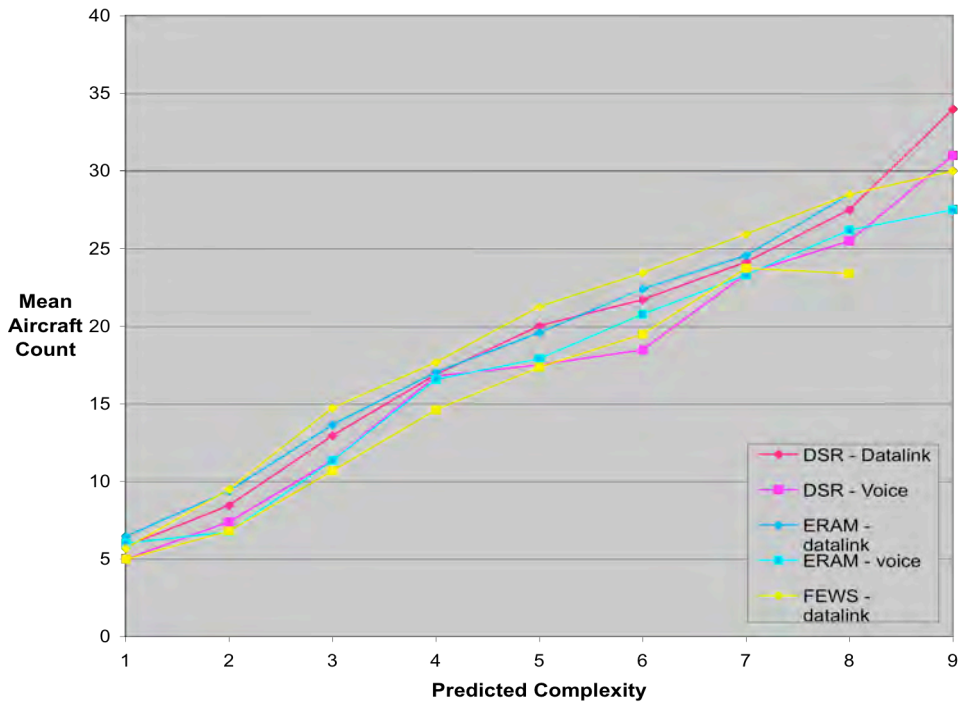
Fig. 7. Average aircraft count by complexity as a function of workstation and communication conditions.
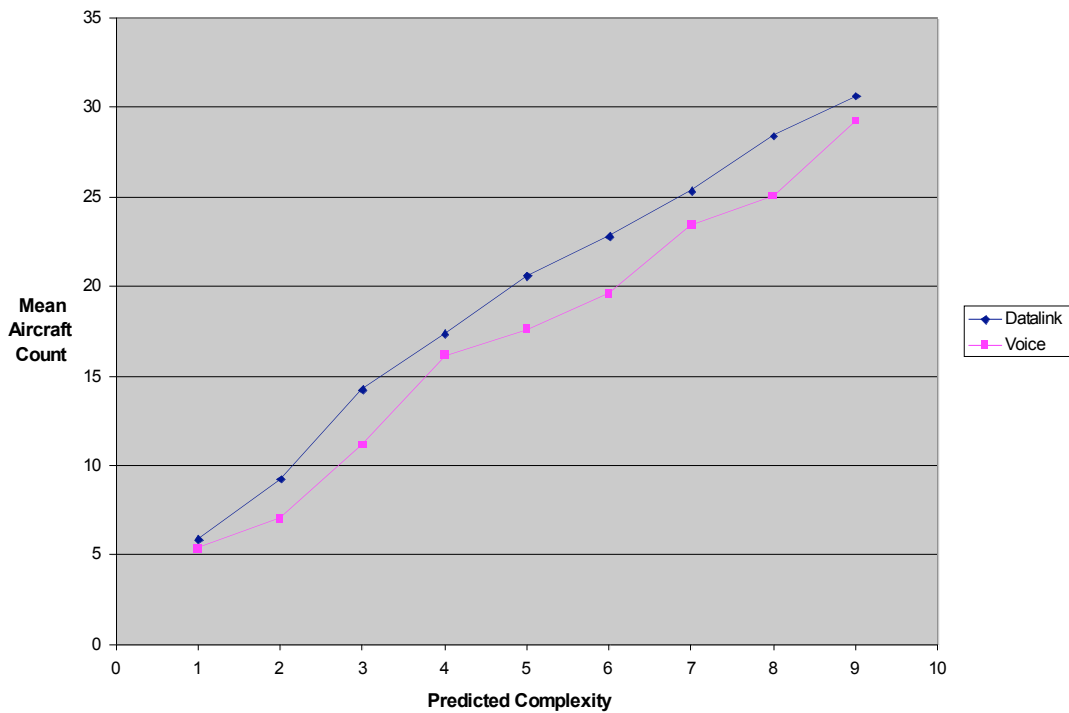

Fig 8. Average aircraft count by complexity as a function of communication condition.

4.6 Losses of Separation and Maximum Manageable Capacity

The study considered a loss of separation to have occurred when aircraft were separated by less than 3 nm horizontally and 1,000 feet vertically. In order to identify a safe complexity range, losses of separations were examined.

Fig. 9 shows the number of separation losses associated with the complexity rating. Controllers were instructed in the HITL study that a workload rating of 6 indicated that they were about to enter an area where it was more likely that they might make a mistake.

The data confirm that many of the separation losses occurred when the traffic was complex or very complex (values 5 and higher). However, Fig. 9 also shows two cases of separation losses when complexity values were 4. This substantiates the utility of a complexity metric and indicates that at higher complexities the chances of separation losses increase. However, results must be interpreted with caution, since this was a simulation of generic airspace and new displays and technologies were used. Although the controllers received training, a lack of

familiarity and exposure may also have contributed to this result [8].

Fig. 10 shows the minimum, average, and maximum number of aircraft that were managed at various complexity levels. This figure clearly indicates that although at times the number of aircraft was below 21, it was close to 21 at complexity range 1-3 which was error free. At higher complexity values, the maximum number of aircraft managed was higher, however, at these complexities, the chance of operational errors also increased as seen in Fig. 9.

Fig. 10 shows a range of aircraft managed by controllers at different complexity levels. It indicates that at lower complexity situations (i.e., complexity range of 1 to just under 4), the range of aircraft was just under 24 corresponding to maximum aircraft count at complexity 4. This range shows a conservative estimate of aircraft count, as there were no errors in this complexity range. At higher complexities, there were losses of separation but they could be partially due to inadequate training related to new displays, procedures, separation standards, and automation.
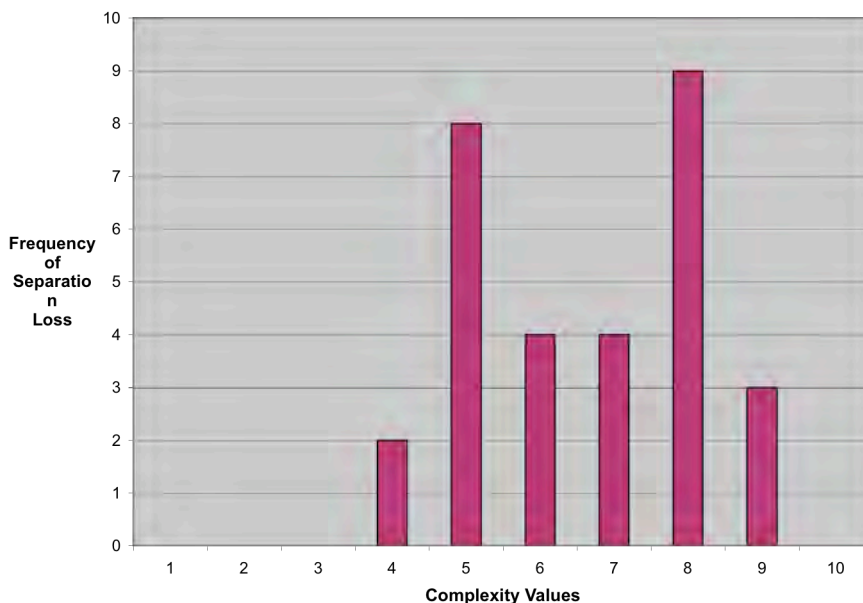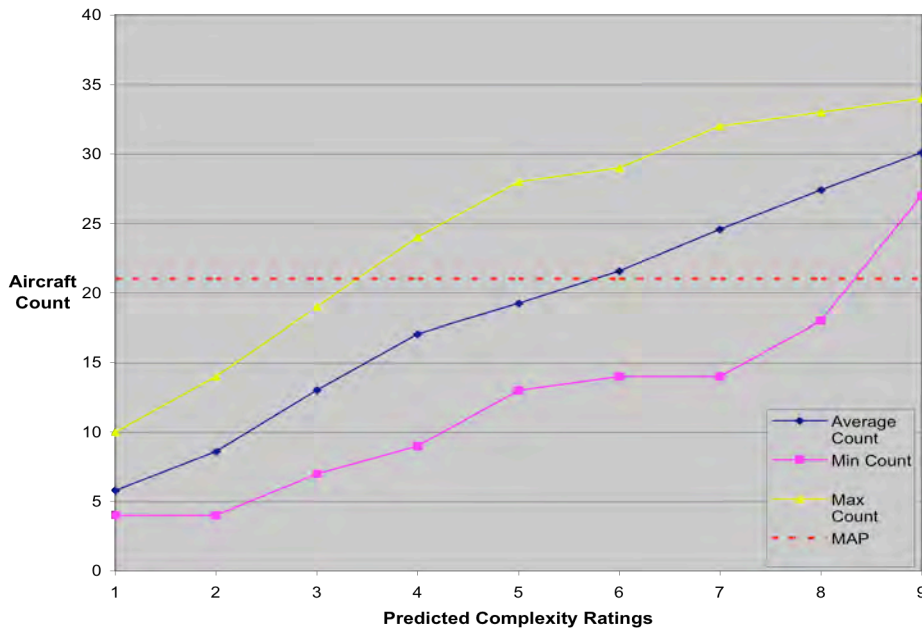

Fig. 9. Complexity and separation loss counts.

Fig. 10. Range of controlled aircraft at various complexity levels.

The operational use of the MAP value of course takes into account that controllers may be able to support instantaneous peaks of traffic that are close to the MAP value. The sustainable aircraft count is usually lower than the MAP value. In this simulation, controllers tried to operate traffic as long as they felt they could handle the workload safely. Hence, the most conservative analysis based on Figures 9 and 10 indicates that the sector capacities, based on complexity, will vary based on individual differences. At times, it will result in increased sector capacity where as at other times it may result in decreased sector capacity as compared with the MAP value.

5   Conclusions

The complexity-based model represented workload better than the aircraft count-based model. The MAP value of the sector was 21. The controllers were able to control more aircraft than the sector's MAP value in many cases. Although the number of aircraft they managed exceeded the MAP value, it also led to losses of separation at higher complexity levels. However, within a complexity range of 1 to just under 4, no losses of separation were observed. The corresponding average number of aircraft controlled within that complexity range was about 24. This indicates that sector capacities will vary if the notion of complexity was used to determine capacity limits. The complexity range that did not contain any instances of loss of separation likely corresponds with levels of traffic that controllers can sustain for longer periods of time. The results should be interpreted with caution as they are based only on analysis of a generic sector.

6   References

[1] Federal Aviation Administration. Order JO 7210.3V. Facility Operation and Administration, Chapter 17, Section 7. Monitor Alert Parameter, 2008.

[2] Arad, B., The Control Load and Sector Design, Journal of Air Traffic Control, pp 13-31, 1964.

[3] Mogford, R., Guttman., S., Morrow, S., and Kopardekar P., The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature,   DOT/FAA/CT-TN-95/22,   FAA Technical Center, Atlantic City, 1995.

[4] Chatterji, G., and Sridhar, B, Measures for Air Traffic   Controller   Workload   Prediction, *Proceedings of the First AIAA Aircraft Technology, Integration, and Operations Forum*, Los Angeles, CA, 2001.

[5] Kopardekar, P., and Magyarits, S., Measurement and Prediction of Airspace Complexity, *5th*

US/Europe Air Traffic Management Seminar, Budapest, Hungary, 2003.

[6] Kopardekar, P., Schwartz, A., Magyarits, S., and Rhodes, J., Airspace Complexity Measurement: An Air Traffic Control Simulation Analysis, *7th US/Europe Air Traffic Management Seminar*, Barcelona, Spain, 2007.

[7] Guttman, J.A. & Stein, E.S. (1997). En route generic airspace evaluation (DOT/FAA/CT-TN97/7). Atlantic City International Airport: Federal Aviation Administration William J. Hughes Technical Center.

[8] Schroeder, D., Bailey, L., Pounds, J., and Manning, C., A Human Factors Review of the Operational Error Literature, DOT/FAA/AM-06/21, Civil Aerospace Medical Institute, Oklahoma City, OK, 2006.

Copyright Statement