



**Federal Aviation
Administration**

DOT/FAA/AM-06/29
Office of Aerospace Medicine
Washington, DC 20591

Relationship of Sector Activity and Sector Complexity to Air Traffic Controller Taskload

Carol A. Manning
Elaine M. Pfeleiderer
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

December 2006

Final Report

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents thereof.

This publication and all Office of Aerospace Medicine technical reports are available in full-text from the Civil Aerospace Medical Institute's publications

Web site:

www.faa.gov/library/reports/medical/oamtechreports/index.cfm

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-06/29		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Relationship of Sector Activity and Sector Complexity to Air Traffic Controller Taskload				5. Report Date December 2006	
				6. Performing Organization Code	
7. Author(s) Manning CA, Pfleiderer EM				8. Performing Organization Report No.	
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No.	
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591				13. Type of Report and Period Covered	
				14. Sponsoring Agency Code	
15. Supplemental Notes Work was accomplished under approved task AM-HRR-524.					
16. Abstract This study compared the relative effectiveness of two constructs, sector activity and sector complexity, in predicting air traffic controller taskload. Sector activity was defined as the activity associated with aircraft moving through the sector and was measured by counting the number of aircraft under the control of the sector during a traffic sample. Sector complexity describes a set of factors presumed to affect the difficulty experienced by a controller when controlling traffic. Sector complexity was measured in two ways. The first measure of complexity was a subjective rating made by supervisors and controllers to describe the complexity associated with specific traffic samples. The second was a composite variable that included measures reflecting several of the complexity variables found in the literature. Taskload was defined as controller activity and was measured by counting the number of data entries made by a controller during a traffic sample. The results appear to suggest that our hypothesis, that sector activity predicted controller taskload better than sector complexity, was incorrect. However, interpretation of these results depended on consideration of what each of the variables measured. The Complexity Rating predicted controller activity better than the number of aircraft alone, but the Complexity Value (based on a set of variables identified through previous research) did not contribute at all to that prediction. Additional analyses suggested that the Complexity Rating measured something very different than the Complexity Value. We believe that the Complexity Ratings estimated the workload that observers believed the controller at the sector experienced instead of the complexity of the situation. On the other hand, the complexity measures used here did not appear to be not good measures of the construct. This may have occurred because the measures used in this study had limited variability or because they were not very good measures of the construct even though they were derived from factors identified in the literature as contributing to sector complexity. While we expected that the number of aircraft alone might be sufficient to predict controller activity/ taskload, the results suggested that measuring both controller activity and extracting measures from other routinely recorded data might be necessary to develop more objective staffing standards used to determine how many controllers are needed to provide ATC services to individual facilities.					
17. Key Words Air Traffic Control, Sector Activity, Sector Complexity, ATC Taskload				18. Distribution Statement Document is available to the public through the Defense Technical Information Center, Ft. Belvoir, VA 22060; and the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 12	22. Price

RELATIONSHIP OF SECTOR ACTIVITY AND SECTOR COMPLEXITY TO AIR TRAFFIC CONTROLLER TASKLOAD

Introduction

The purpose of this paper is to compare the relative effectiveness of measures of two constructs, sector activity and sector complexity, in predicting air traffic controller taskload. In spite of research that suggests that the number of aircraft alone is insufficient to describe a controller's workload, we argue that it is a good predictor of controller taskload. In fact, we believe that the number of aircraft is as good, if not better, a predictor of controller taskload than are a set of measures of sector complexity developed to predict when sectors will become too difficult for one controller to work. It is useful to be able to predict controller taskload because certain measures, such as FAA's staffing standards (used to determine the number of controllers allowed to be on staff at a facility), are based on observations of the number and time a controller spends performing ATC-related tasks. If those observations can be replaced by more efficient or objective measurements of controller taskload, then the staffing standards could be updated more frequently.

Before addressing this issue, it is first necessary to define terminology. In this study, the criterion is air traffic controller taskload, defined here as controller activity, and measured by counting the number of data entries made during a traffic sample. The predictors in this study are sector activity, defined as the activity associated with aircraft moving through the sector, and measured by counting the number of aircraft under the control of the sector during a traffic sample. Sector complexity is a more difficult construct to measure than is sector activity. Sector complexity (also known as "air traffic complexity," "cognitive complexity," and sometimes "dynamic density") is a term used to describe a set of factors presumed to affect the difficulty experienced by a controller when controlling traffic. The sector complexity construct is proposed to describe the important factors associated with a sector (besides simply the presence of aircraft) that can make the job of controlling aircraft more difficult. Complexity factors, as defined by different authors, often include such variables as the presence of climbing or descending aircraft, aircraft mix (different types of aircraft having different performance characteristics), special use airspace activity, and presence of severe weather.

Several researchers (Mogford, Murphy, & Gutman, 1994; Wyndemere, 1996; Laudeman, Shelden, Branstrom, & Brasil, 1998; Sridhar, Sheth, & Grabbe, 1998; Christien & Benkouar, 2003; Histon, Aigoin, Delahaye, Hansman, & Puechmorel, 2001; and Majumdar & Ochieng, 2000) have proposed different ways to measure sector complexity. Hilburn (2004) listed a set of 108 complexity factors comprising categories that encompassed most of the complexity factors identified in the literature. Most of these measures describe factors such as the variability among the aircraft in a sector (e.g., number of aircraft, percent of climbing/descending aircraft, aircraft mix, presence of military aircraft) variability among the physical characteristics of a sector (e.g., size; shape; number, direction, and intersection of flight paths and levels; number of shelves; and presence and activity of special use airspace), types of procedures used (e.g., required procedures, amount of coordination required, number of facilities with which interactions occur, complex routings, sequencing, and spacing), specific traffic circumstances (e.g., number and distance of conflicts, distance between aircraft) the extent to which operations are affected by winds and severe weather, and adequacy of radio or radar coverage.

For this study, sector complexity was defined two ways. The first measure of complexity was a subjective rating made by supervisors and controllers to describe the complexity associated with specific traffic samples (derived from data collected by Kopardekar & Magyrits, 2003). The second was a composite variable that included measures reflecting several of the complexity variables found in the literature. The measures that were combined to form the composite complexity variable were: visual clustering (2-dimensional distances between aircraft), special use airspace (SUA) activity, presence of severe weather, frequency congestion, amount of coordination required, percentage of climbing/descending aircraft, and aircraft mix.

The purpose of this study was to compare the effectiveness of the sector activity and sector complexity constructs to see which was a better predictor of controller taskload. An argument is often made that controller taskload is influenced by variables other than the number of aircraft in the sector. For example, Laudeman et al.

(1998) found that, while the number of aircraft alone accounted for 33% of the variance in controller activity ($r = .57$), the prediction (r^2) was increased to 55% ($r = .74$) when using their Dynamic Density factors. However, other research has found a strong relationship between the number of aircraft and taskload measures. Manning, Mills, Fox, Pfeleiderer, and Mogilka (2001) and Manning, Mills, Fox, Pfeleiderer, and Mogilka, (2002) found that the number of aircraft in a sector during a traffic sample loaded on the same principal component as several measures of controller activity, while variables related to sector complexity were more closely related to a different orthogonal component. These results suggest that the number of aircraft may be a better predictor of taskload than complexity, at least when the variables are measured using routinely recorded ATC data.

We anticipated that sector activity would be a better predictor of controller taskload than sector complexity and that adding sector complexity to the regression equation would not contribute anything unique to the prediction of taskload over that contributed by sector activity alone.

Method

Data

The traffic samples came from a study conducted by Kopardekar and Magyrits (2003) to compare alternative dynamic density metrics. The data were extracted from ninety 6-minute traffic samples of routinely recorded System Analysis Recording (SAR) data collected from the Atlanta and Ft. Worth Air Route Traffic Control Centers, totaling 180 samples. Besides the SAR data, pilot-controller communications recorded during the same traffic samples were also analyzed. Sixteen traffic samples with missing communication files were excluded from the analysis, leaving 164 samples.

In addition, real-time controller and supervisor ratings of complexity were made for the same traffic samples at 2-minute intervals. The rating data were first averaged across raters, then the composite ratings for three 2-minute intervals were averaged together to provide one rating for each 6-minute interval.

The SAR files were first processed using the FAA's Data Analysis and Reduction Tool (DART; Federal Aviation Administration, 1993), which produced output files that were then processed using the NAS Data Management System (NDMS) and POWER (Mills, Pfeleiderer, & Manning, 2002). The result was a set of more than 20 variables describing controller and aircraft activity. The voice tapes containing pilot/controller communications for the same traffic samples were transcribed, numbers

and timing of transmissions were recorded, and content of the transmissions was coded.

Computation of variables

Taskload. Taskload was estimated by measuring observable controller activity. Controller activity is defined here as the number of data entries (number of commands, not individual keystrokes entered) made by a controller during the course of a traffic sample. Data entries are recorded at both the Radar (R) and Radar Associate (RA) positions at a sector, regardless of whether one or two controllers are physically present (because some entries can only be made at each position). The variable used in this study to describe controller activity was the sum of the R and RA data entries made at the sector during the traffic sample.

Sector activity. Sector activity is related to the number of aircraft controlled by the R controller during the traffic sample. The number of aircraft might be characterized as either the total number of aircraft controlled or the maximum number of aircraft controlled at any one time during the traffic sample. We initially considered both the total number of aircraft and the maximum number of aircraft to represent sector activity. Analyses reported below suggested it was not necessary to include both variables in subsequent analyses, so we retained only the total number of aircraft controlled during a traffic sample.

Sector complexity. Several studies have identified measures contributing to sector complexity. Most of these include characteristics associated with aircraft movement, such as the amount of climbing or descending aircraft and aircraft mix, or circumstances associated with specific traffic situations including military/special use airspace activity, presence of severe weather, amount of coordination required, or frequency congestion.

In what way might these variables affect sector complexity? In general, an increased amount of these factors would be expected to increase the difficulty of dealing with specific traffic situations, create distractions, or reduce a controller's ability to be flexible when controlling aircraft. More specifically, when aircraft climb or descend, it may be difficult for controllers to estimate how long it will take them to reach or leave an altitude. It may also be difficult to project whether two aircraft will conflict laterally if one is simultaneously moving both laterally and vertically.

Aircraft mix may contribute to complexity because controllers must take differences in aircraft performance characteristics into account when determining how to sequence a line of aircraft into an airport and when deciding which aircraft to climb or descend to ensure separation is maintained. SUA activity limits the airspace

available, thus limiting the possible choices of solutions to sequence aircraft or avoid conflicts. Similarly, the presence of severe weather can reduce available airspace, thereby constraining control options. Severe weather may also create new sector “choke points,” requiring controllers to watch for conflicts in unusual locations and use non-typical routings for aircraft.

The amount of coordination required can increase sector complexity because coordination requires controllers to focus on communication rather than scanning the aircraft in their sector or formulating plans to sequence aircraft or maintain separation. Frequency congestion is similarly distracting but is also a problem because it can reduce the ability of a controller to contact a pilot to issue a clearance required to avoid a conflict.

Table 1 lists the complexity variables described above and provides a brief definition for each. *Complexity Ratings* were subjective judgments made by controllers and supervisors who watched the traffic samples using the SATORI re-creation tool (Rodgers & Duke, 1993). The ratings were made on a 7-point scale, were obtained every 2 minutes, and were averaged over time and across controllers. *SUA Activity* and *Weather* were based on observations made by controller and supervisor subject matter experts (SMEs) about whether special use airspace was active or severe weather was present during each traffic sample. *Climbing and Descending Aircraft* was the percentage of aircraft in the traffic sample found to be climbing or descending, rather than in level flight, as determined from their recorded flight profiles.

Three variables were identified that could potentially represent *Coordination Required*. These were the number of pointouts made by either the R or RA controller, the number of off-frequency messages made in each traffic sample, and the average amount

of time required for off-frequency messages. Because the correlation between the number of off-frequency messages and the time required to make them was .95, only the number of pointouts and the time required for off-frequency messages were retained for analysis.

Visual Clustering described the number of aircraft that came within x lateral miles of another aircraft, regardless of the amount of altitude separation. This measure was based on Stein’s (1985) measure of “local density.” The variable was termed *Visual Clustering* because high values of this variable would be associated with aircraft that appeared to cluster together when looking at the 2-dimensional radar display even though they might be separated by altitude. In this study, the lateral separation value used was 10 nautical miles.

Frequency Congestion was measured as the amount of time either the R controller or a pilot spoke on the radio and was obtained from the voice communication recordings. Communication time may also be considered a measure of taskload, but because Mogford et al. (1994) identified frequency congestion as a complexity factor, we used it in that manner for this study.

Aircraft Mix was a number indicating the difference in engine types between all aircraft in a sector (Pfleiderer, 2003). The aircraft mix index was computed by first, assigning aircraft type codes with values ranging from one (for Piston-driven aircraft) to four (for high-performance jets) to all aircraft. Second, a half-matrix of aircraft type differences (the absolute value of the difference between the aircraft type codes for each aircraft pair) was computed. Third, all values of the aircraft type differences in the half matrix were summed. The aircraft mix index was computed for all aircraft present during each radar update (approximately 12-second intervals) and averaged over each 6-minute traffic sample.

Table 1. Sector complexity measures.

Name	Description
Complexity Ratings	Subjective complexity ratings made by controllers and supervisors
SUA	Controller assessment - was special use airspace active?
Weather	Controller assessment - was severe weather present?
Climbing & Descending Aircraft	% climbing, descending aircraft
Coordination Required	# pointouts made, time required for off-frequency messages
Visual Clustering	# aircraft within 10 nm
Frequency Congestion	Amount of transmission time
Aircraft Mix	Measure reflecting difference in engine types

Results

Descriptive statistics

Taskload. The number of R and RA controller data entries was the only criterion variable. An average of 34.2 R and RA controller entries ($SD = 11.7$) were made per traffic sample ($N=164$).

Sector activity. Table 2 shows descriptive statistics and correlations for the two variables proposed to represent sector activity, the total number of aircraft and the maximum number of aircraft controlled simultaneously. Both variables were computed for each of the 6-minute traffic samples analyzed. The variables had a statistically significant correlation of .89 and were, thus, considered nearly equivalent. For this reason, only the total number of aircraft controlled was retained for analysis.

Sector complexity. Table 3 shows descriptive statistics and correlations for the sector complexity variables. SUA activity and weather were both bivariate distributions (with

values of 0=not present and 1=present.) For about 75% of the traffic samples, SUA activity was equal to 0. For about 60% of the traffic samples, presence of significant weather was equal to 0. Distributions of the number of pointouts, number of off-frequency transmissions, off-frequency transmission time, and aircraft mix were positively skewed. For about half the traffic samples, each of these variables had values equal to 0.

Complexity ratings were significantly correlated with five of the eight complexity variables. Visual clustering was correlated with five of the other sector complexity variables. SUA activity and the proportion of climbing and descending aircraft were significantly correlated with four other variables.

Regression analysis

To reduce the number of variables used for analysis, we often compute a Principal Components Analysis to identify a set of components that account for the

Table 2. Descriptive statistics and correlations between sector activity variables ($N=164$).

Name	Average	Std. Dev	Correlations
Total # aircraft	15.0	4.0	
Maximum # aircraft controlled simultaneously	11.3	3.2	.89**

**Significant at $p < .01$.

Table 3. Descriptive statistics and correlations between sector complexity variables ($N=164$).

Name	Avg	SD	Correlations							
			CR	SUA	Wx	%CD	PO	OFT	VC	FC
Complexity ratings	3.4	1.1								
SUA	0.2	0.4	-.06							
Weather	0.4	0.5	.24*	-.21**						
% of climbing & descending aircraft	0.5	0.2	-.26*	-.28**	-.04					
Coordination required (POs)	1.4	1.9	.09	-.19*	.23**	-.01				
Off freq message time	15.7	20.8	-.39**	.02	.08	.07	.02			
Visual clustering	6.2	3.0	.45	.17	-.05	-.06	-.03	-.24		
Frequency congestion	216.1	55.7	.57**	.08	.03	.41**	.03	-.09	.41**	
Aircraft mix	11.3	15.3	-.13	-.05	-.12	.23	-.05	.01	.23	.06

* $p < .05$. ** $p < .01$.

variability in a set of variables and use the component scores to predict the criterion measure. However, in this analysis, the sector complexity variables were assumed to be of equal weight, and a higher value of each variable was presumed to be associated with a higher amount of complexity. Consequently, z-scores were computed for all sector complexity variables (except the Complexity Rating) and were summed to compute a composite complexity measure for each traffic sample. We then used this composite complexity variable, called “Complexity Value,” in subsequent analyses. The Complexity Rating was not standardized, but was instead retained as a separate variable.

Table 4 shows the correlations between the variables entered into the regression analysis, number of R and RA data entries, total number of aircraft, the Complexity Rating, and the Complexity Value. All zero-order correlations were significant except the correlation between the Complexity Value and the total number of aircraft and the correlation between the Complexity Value and the number of R and RA data entries. The correlation between the Complexity Value and the Complexity Rating, while statistically significant, was lower than the correlations between the other variables. Notably, the zero-order correlation between the number of aircraft and the average complexity ratings was very high ($r = .66$).

A set of analyses was performed to assess the effectiveness of alternative multiple regression models in predicting controller activity. We used a method that allowed us to compare specific regression models instead of an analysis such as stepwise linear regression because we wanted to assess the relative contribution of specific variables to the model rather than simply those variables that made

statistically significant contributions, such as would result when conducting a multiple regression analysis.

Table 5 shows the results of these analyses. Row 1 shows the multiple correlation of the full model containing all three predictor variables (Number of Aircraft, Complexity Rating, and Complexity Value) with the criterion variable (number of R and RA controller data entries). The multiple correlation of the full model with the number of R and RA controller data entries was $R = .62$, accounting for about 39% of the variance in controller activity. Succeeding lines show multiple correlations between alternative (reduced) regression models containing fewer than the total number of predictors.

The column containing F for the test of R^2 change compares the relative effectiveness of a reduced model with the effectiveness of the full model in predicting the number of R and RA data entries. If the probability is greater than .05 that the change in R^2 between the two models is significantly different from zero, then the reduced model is considered to be as effective as (i.e., no different than) the full model in predicting controller activity. On the other hand, if the probability is less than or equal to .05 that the change in R^2 between the two models is significantly different from zero, then the reduced model is not considered to be as effective as the full model in predicting controller activity. The goal of the analysis is to identify a reduced model that contains as few predictors as possible but accounts for a high enough percentage of the variance in the dependent variable to be considered equivalent to the full model. In this way, we will assess the relationship between each specific predictor and the criterion measure of controller taskload.

Table 4. Descriptive statistics and correlations between variables included in multiple regression model ($N=164$).

Name	Avg	SD	Correlations			
			# R&RA data entries	N aircraft	Complexity Rating	Complexity Value
R & RA data entries	41.1	13.7	1.0			
N aircraft	15.0	4.0	.53**	1.0		
Complexity Rating	3.45	1.11	.58**	.66**	1.0	
Complexity Value	0.0	3.0	.01	.11	.17*	1.0

* $p < .05$. ** $p < .01$.

Table 5. Results of analyses comparing alternative multiple regression models prediction of controller activity.

Regression model	R	R^2	R^2 change	F for test of R^2 change	df	p
1. Full model containing number of aircraft, Complexity Rating, and Complexity Value	0.62	0.39	N/A	N/A		
Individual predictors						
2. Model containing only number of aircraft	0.53	0.29	0.10	13.46	2, 160	.000
3. Model containing only the Complexity Rating	0.58	0.34	0.05	6.31	2, 160	.002
4. Model containing only Complexity Value	0.01	0.00	0.39	50.72	2, 160	.000
Predictor pairs						
5. Model containing n aircraft and Complexity Rating	0.62	0.38	0.01*	1.95	1, 160	.165**
6. Model containing N aircraft and Complexity Value	0.54	0.29	0.10	26.37	1,160	.000
7. Model containing Complexity Rating and Complexity Value	0.59	0.35	0.04	10.78	1,160	.001

* Difference in R^2 was greater than 0 due to rounding.

** Indicates reduced models that predicted controller data entries as well as the full model.

The analysis of six reduced models is shown in Table 5 (see rows 2-7). The first group of analyses (rows 2-4) compared three reduced models consisting of individual predictor variables with the full model. The second group of analyses (rows 5-7) compared three different reduced models consisting of pairs of the different predictor variables with the full model.

As an example, Row 2 compared a reduced model containing only the number of aircraft as a predictor with the full model. The model containing the number of aircraft had an R^2 of .29, compared with the full model's R^2 of .39. The F computed to assess the R^2 change of .10 had a value of 13.46. The probability was less than .001 that the change in R^2 was greater than zero. The reduced model containing only the number of aircraft was significantly different than the full model in predicting controller activity and, thus, was not as effective as the full model.

The only model that predicted controller activity as well as the full model is shown on Row 5. The reduced model, containing both the number of aircraft and the Complexity Rating, had an R^2 of .62, the same as the full model. The F computed to compare the R^2 change of .01 (which was greater than 0 due to rounding) had a value of 1.95, and the probability was less than .17 that the change in R^2 was greater than 0. Thus, the reduced model containing both the number of aircraft and the Complexity Rating predicted the number of R and RA data entries as well as the full model.

The model containing only the Complexity Value contributed virtually nothing to the prediction of R and RA controller data entries. Furthermore, when the variable was entered into a model with either of the other predictors, it also added nothing to the prediction of the criterion measure. Note that while both the number of aircraft and the Complexity Rating had significant zero-order correlations with the criterion measure, neither, in isolation, was sufficient to predict the criterion as well as the full model.

Principal Components Analysis

A Principal Components Analysis was conducted using the four variables to assess their interrelationships. Based on previous research (Manning et al., 2002; Pfeleiderer, 2005), we expected that the total number of aircraft and the number of data entries would load on a single *Activity* component. We also expected that, because they purportedly measured the same construct, the two complexity variables would load on a second *Complexity* component, which should be different from *Activity*.

Two components were extracted. Table 6 shows the factor loadings, rotated with the Varimax rotation method. The first component extracted had an eigenvalue of 2.21 (accounting for 55 % of the variance in the data), while the second had an eigenvalue of 1.00 (accounting for 25% of the variance in the data).

Table 6. Rotated component matrix.

Variable name	Components	
	Component 1: 'Activity'	Component 2: 'Complexity'
R & RA data entries	.83	
N aircraft	.85	
Complexity Rating	.87	
Complexity Value		.99

Note: Component loadings < .3 were eliminated.

The results were only partly consistent with our expectations. As expected, Component 1 was clearly related to controller and sector activity. However, three variables instead of two--the Complexity Rating, R and RA data entries, and the Number of Aircraft--had high and equivalent correlations with this component. The second component was defined exclusively by the Complexity Value variable. The Complexity Value did not correlate significantly with the *Activity* component, nor did any of the other three variables correlate significantly with the *Complexity* component.

Conclusions

The results appear to suggest that our hypothesis that sector activity would predict controller taskload better than sector complexity was incorrect. The regression analyses showed that only a model containing both the number of aircraft and the Complexity Rating was equivalent to a full model containing all three variables when predicting controller taskload.

However, when interpreting these results, the reader must consider what each of the variables measured. We found that the Complexity Rating (based on SMEs' subjective judgments about the complexity of the traffic samples) predicted controller activity better than the number of aircraft alone, but the Complexity Value (based on a set of variables identified through previous research as contributing to sector complexity) did not contribute at all to that prediction. A subsequent Principal Components Analysis revealed that the number of R and RA data entries, the number of aircraft, and Complexity Ratings loaded equally well on a component best described as *Activity* while the Complexity Value was the only variable that loaded on a separate component that we called *Complexity*. This result suggests that the Complexity Rating variable measured something very different than the Complexity Value.

The Complexity Rating was measured on a 7-point Likert-type scale where 1 represented low complexity and 7 represented high complexity. The study for which

these data were collected did not define these terms and did not collect SME feedback about how they interpreted complexity. Thus, it is not clear what the controllers were estimating when they provided the Complexity Ratings. Perhaps the Complexity Ratings were instead an estimate of the workload that observers believed the controller at the sector experienced. This interpretation is supported by the strong relationship between the Complexity Rating, data entries, and number of aircraft, and the corresponding lack of a relationship between Complexity Rating and Complexity Value. Similar relationships were observed between data entries, number of aircraft, and a real-time subjective workload assessment (when measured using the Air Traffic Workload Input Technique during Manning et al., 2002) when all variables loaded on a similar *Activity* component.

These results suggest that activity and workload are very similar. First, the controllers did not appear to rate Complexity because the PCA showed that their ratings loaded on an *Activity* component rather than a *Complexity* component. Second, the Complexity Value, as measured here, was not very closely related to either sector activity or controller taskload because it loaded on a completely different component than did those variables.

Perhaps the complexity measures used here were not good measures of the construct. Even though the individual measures used in this study were derived from factors identified by numerous researchers as contributing to sector complexity, certain distributional problems may have limited their effectiveness. Distributions of the individual variables included in the composite did not have much variability, which may have indicated that the traffic samples used were not very complex. The mean of the complexity ratings was 3.4, less than the halfway point on the 7-point Likert scale. Only 29% of the individual complexity ratings exceeded the mid-point. We often find that samples of real traffic are of relatively low workload (as seen in Manning et al., 2002). Perhaps complexity metrics should be tested in simulations (like those conducted by Lee, 2005) that increase traffic and other variables beyond normal levels to produce higher

values of complexity. Alternatively, these variables may be insufficient, and different variables should be developed to represent the complexity of traffic in the real world.

The third part of this discussion deals with the relationship between taskload, workload, and aircraft activity. If we assume that the complexity ratings in this study actually measured workload, it suggests that 1) activity and workload are very similar and yet 2) activity and workload are better predictors of taskload when paired together, even though both are fairly good predictors in isolation.

Perhaps we should be asking: What are we trying to predict, and how should we measure it? One set of studies attempts to predict workload with the goal of determining whether new systems or procedures will unduly increase it. Another set of studies develops complexity measures for the purpose of predicting sector activity at some time in the near future (such as 20 minutes in advance), although the present study suggests that there is not much of a relationship between sector complexity variables and current sector activity.

Studies that attempt to predict taskload do not occur often in the literature. However, taskload prediction could be useful for setting staffing standards that are used to determine how many controllers are needed to provide ATC services to individual facilities. Previous efforts to set staffing standards were based on time-and-motion studies, in which observers recorded how long it takes controllers to perform observable activities. Software such as POWER (Mills et al., 2002) would facilitate obtaining such measures through analysis of routinely recorded data, not requiring special observational studies (that might influence controller behavior) to be conducted.

We proposed in this paper that the number of aircraft alone might be sufficient to predict controller activity/taskload, but the results suggest that this is not accurate. However, perhaps measuring controller activity (e.g., counting data entries) and extracting measures from other routinely recorded data could provide sufficient information to develop more sophisticated, or at least more objective, staffing standards.

It should be noted that workload was not explicitly measured in this study. While it appears that the SMEs did not rate sector complexity, we cannot be certain whether they rated workload instead. Consequently, it is not appropriate for us to make specific statements about how these results about the prediction of taskload may relate to the prediction of workload. Research on the relationship between those variables is still needed.

References

- Christien, R. & Benkouar, A. (2003, June). Air traffic complexity indicators and ATC sectors classification. In *Proceedings of 5th Eurocontrol/FAA Air Traffic Management R&D Seminar*, Budapest, Hungary.
- Federal Aviation Administration. (1993). *Multiple Virtual Storage (MVS); User's manual; Data Analysis and Reduction Tool (DART)*. (NASP-9247-PO2). Washington, DC: Author.
- Hilburn, B. (2004, March). *Cognitive complexity in air traffic control: A literature review*. (EEC Note No. 04/04). Brétigny-sur-Orge: France: Eurocontrol Experimental Centre.
- Histon, J.M., Aigoïn, G., Delahaye, D., Hansman, R.J., & Puechmorel, S. (2001, December). Introducing structural considerations into complexity metrics. In *Proceedings of 4th Eurocontrol/FAA Air Traffic Management R&D Seminar*, Santa Fe, NM.
- Kopardekar, P. & Magyrits, S. (2003, June). Measurement and prediction of dynamic density. In *Proceedings of 5th Eurocontrol/FAA Air Traffic Management R&D Seminar*, Budapest, Hungary.
- Laudeman, I., Shelden, S., Branstrom, R., & Brasil, C. (1998). *Dynamic density: An air traffic management metric*. Report NASA/TM-1998-112226. Ames Research Center, Moffett Field, California: NASA.
- Lee, P. (2005, October). A nonlinear relationship between controller workload and traffic count. In *Proceedings of Human Factors and Ergonomics Society 49th Annual Meeting*, Orlando, FL.
- Majumdar, A. & Ochieng, W.Y. (2000). *The factors affecting air traffic controller workload: A multivariate analysis based upon simulation modelling of controller workload*. Centre for Transport Studies Department of Civil and Environmental Engineering Imperial College of Science, Technology and Medicine. London, United Kingdom.
- Manning, C., Mills, S., Fox, C., Pfleiderer, E., & Mogilka, H. (2001, July). *Investigating the validity of Performance and Objective Workload Evaluation Research (POWER)*. (DOT/FAA/AM-01/10). Washington, DC: FAA Office of Aerospace Medicine.

- Manning, C., Mills, S., Fox, C., Pfliegerer, E., & Mogilka, H. (2002). The relationship between air traffic control communication events and measures of controller taskload and workload. *Air Traffic Control Quarterly*, 10(2), 69-83.
- Mills, S.H., Pfliegerer, E.M., & Manning, C.A. (2002, February). *POWER: Objective activity and taskload assessment in en route air traffic control*. (DOT/FAA/AM-02/2). Washington, DC: FAA Office of Aerospace Medicine.
- Mogford, R., Murphy, E., & Gutman, J. (1994). Using knowledge exploration tools to study airspace complexity in air traffic control. *The International Journal of Aviation Psychology*, 4(1), 29-45.
- Pfliegerer, E.M. (2003, May). *Development of an empirically-based index of aircraft mix*. (DOT/FAA/AM-03/8). Washington, DC: FAA Office of Aerospace Medicine.
- Pfliegerer, E.M. (2005, August). *Relationship of the aircraft mix index with Performance and Objective Workload Evaluation Research (POWER) measures and controllers' subjective complexity ratings*. (DOT/FAA/AM-05/16). Washington, DC: FAA Office of Aerospace Medicine.
- Rodgers, M.D., & Duke, D.A. (1993). SATORI: Situation Assessment Through Re-creation of Incidents. *The Journal of Air Traffic Control*, 35(4), 10-14.
- Sridhar, B., Sheth, K.S., & Grabbe, S. (1998, December). Airspace complexity and its application in air traffic management. In *Proceedings of 2nd Eurocontrol/FAA Air Traffic Management R&D Seminar*, Orlando, FL.
- Stein, E.S. (1985). *Air traffic controller workload: An examination of workload probe*. (DOT/FAA/CT-TN84/24). Atlantic City, NJ: Federal Aviation Administration Technical Center.
- Wyndemere (1996). An evaluation of air traffic control complexity (Contract number NAS2-14284). Boulder, CO: Author.

