Wright State University
CORE Scholar

International Symposium on Aviation Psychology - 2007

International Symposium on Aviation Psychology

2007

Using Multi-Level Analysis to Model the Sources of Variability in Workload Within and Between Sectors

Andrew Neal

Martijn Mooij

Scott Bolland

Tania Xiao

Peter Lindsay

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2007

Part of the Other Psychiatry and Psychology Commons

Repository Citation

Neal, A., Mooij, M., Bolland, S., Xiao, T., & Lindsay, P. (2007). Using Multi-Level Analysis to Model the Sources of Variability in Workload Within and Between Sectors. *2007 International Symposium on Aviation Psychology*, 463-468.
https://corescholar.libraries.wright.edu/isap_2007/56

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2007 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.

USING MULTI-LEVEL ANALYSIS TO MODEL THE SOURCES OF VARIABILITY IN WORKLOAD WITHIN AND BETWEEN SECTORS

Andrew Neal, Martijn Mooij, Scott Bolland, Tania Xiao, Peter Lindsay
The University of Queensland
Brisbane, Australia

Christine Boag-Hodgson
Airservices Australia
Canberra, Australia

Air-traffic control workload data was collected from 20 enroute radar sectors in northern Australia. Multi-level analyses were used to model the effects of traffic variables and airspace variables. The effect of aircraft count varied significantly between sectors. Aircraft count had a stronger effect on workload in sectors that typically have a larger proportion of aircraft on descent, and aircraft that are more closely spaced. Implications for the development of predictive workload metrics are discussed.

Introduction

Controller workload represents one of the most significant constraints on the capacity of the air traffic management system (Leiden, Kopardekar, & Green, 2003). Despite over 40 years of research, our ability to predict the level of workload that a controller will experience given a particular flow of traffic within a particular sector remains limited. It is even more difficult to predict how the workload of a controller will change if the airspace is redesigned, or new technologies or procedures are introduced. Given the projected growth in traffic volume over the next decade, and the need to assess and mitigate any risks associated with airspace changes and sector redesign, we need tools that are capable of answering these questions.

Prior Research

There is a long history of workload research in air traffic control (Arad, 1964; Boag, Neal, Loft, & Halford, 2006; Davis, Danaher, & Fischl, 1963; Hurst & Rose, 1978; Jolitz, 1965; Schmidt, 1976). In a recent review of this literature, Loft, Sanderson, Mooij and Neal (in press) identified three broad categories of variables that have been examined: traffic variables; airspace variables; and operational constraints.

Traffic variables describe the distribution of traffic within a sector. These variables are typically computed in real time using radar track data. Aircraft count is the most widely used factor (e.g., Kopardekar & Magyarits, 2003; Manning et al. 2001). Aircraft under the controller’s jurisdiction create workload for the controller, because they require monitoring and intervention. Traffic density has also been widely studied. A wide range of density metrics have been developed assessing factors such as horizontal and vertical proximity (Chatterji & Sridhar, 2001). Aircraft that are in close proximity to each other create workload, because they require additional monitoring beyond that required for routine traffic.

Airspace variables describe the underlying structural properties of the airspace. Examples include the volume of the sector, aspect ratio (ratio of width to length), number of flight levels, and number of crossing altitude profiles. The design of the sector affects the flow of traffic within that sector, and creates differences among sectors in the typical traffic patterns that controllers deal with. For example, the placement of the routes and boundaries will influence the average number of aircraft that controllers have under jurisdiction, and the average proximity of those aircraft.

Operational constraints also shape the flow of traffic within sectors. Examples of operational constraints include active military airspace, weather, and flow restrictions. These factors restrict the airspace that is available for traffic, can affect the flow of traffic through that airspace, and place constraints on the available actions for controllers and pilots.

The Problem

While there is general agreement regarding the types of factors that predict controller workload, there is less agreement concerning the weighting of these factors. The standard way in which researchers have attempted to build workload metrics is to collect subjective workload ratings from controllers while performing their job, or in the simulator. Track and flight plan data is then extracted from the system, and used to calculate a set of predictive metrics. The researcher then uses multiple regression to identify the metrics that best predict the workload ratings, and calculate the weights for each of these factors. The problem is that different studies produce different results.

Part of the reason why results vary across studies is that different research groups use different predictive metrics. Obviously, the results of a regression analysis depend on the variables that are included in the equation. However, we believe that this is not the whole story. Kopardekar and Magyarits (2003) collected data from 36 enroute sectors and found that the same regression model did not fit all sectors equally well. Their results suggest that different factors predict workload in different sectors. This finding is consistent with anecdotal reports by experienced controllers, who claim that aircraft count and traffic density can have different effects in different sectors.

An example that is often cited concerns the difference between the two sectors in Figures 1 and 2. These two sectors are similar in size, and average traffic volume. Keppel/Alma is an enroute sector north of Brisbane, with segregated parallel one-way routes. Daintree/Reef-Low, by contrast, is centered on a major airport (Cairns) with a converging route structure and only the routes to the south of the major airport are segregated one-way tracks. Both are combined low and high altitude sectors and both have a mix of jet and propeller aircraft. Controllers report that due to the converging route structure and the sequencing of aircraft into the major airport, the nature of the factors driving workload differs quite substantially for Daintree/Reef-Low than for Keppel/Alma.

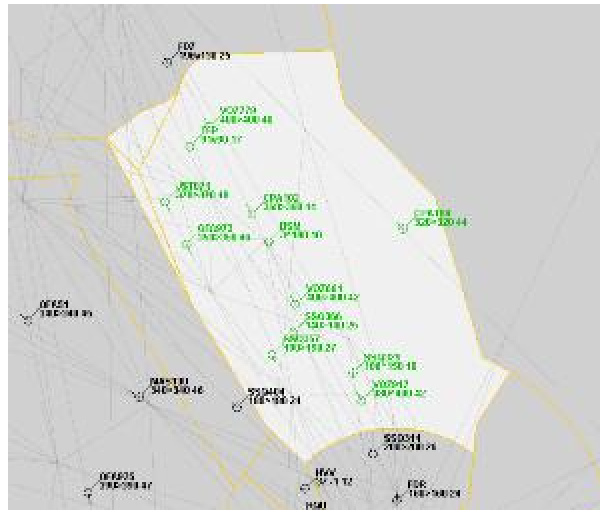

Figure 1. Planview display of Keppel/Alma

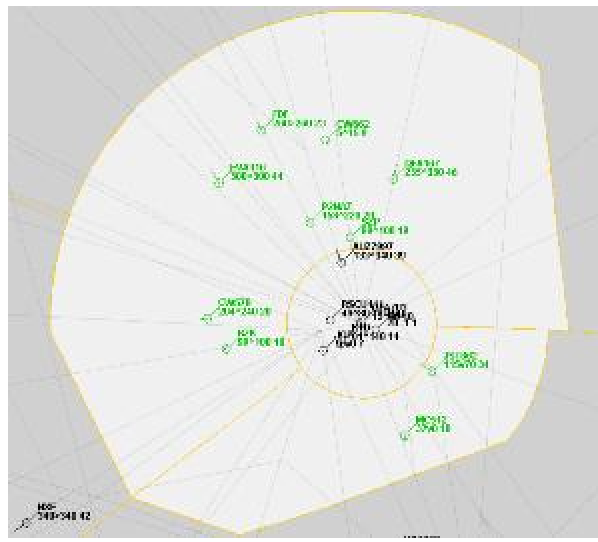

Figure 2. Planview display of Daintree/Reef-Low

Differences among sectors would not be a problem if the structure of the airspace was unchanging. In theory, it would be possible to obtain workload ratings for a representative sample of traffic in all sectors, and empirically derive the best fitting regression model for each sector. Thus, each sector would have a different workload equation. However, there are several reasons why this approach is not feasible. First, it is difficult to obtain enough operational data for each sector. There are practical constraints that limit the amount of data that can be collected in an operational environment. More importantly, however, regression models that have been calibrated on individual sectors cannot be used for sector redesign. If the weightings on the parameters in the model vary across sectors, then modifying the sector by changing the boundaries or route structure should change the parameters in the model.

The aim of the current paper is to introduce a statistical method for solving this problem. The method that we use is random coefficient modeling (Bleise & Ployhart, 2002).

Random Coefficient Modeling

Random coefficient models allow the investigation of phenomena at different levels of analysis. In the current paper, we are interested in variables at two levels of analysis: the *within-sector* level, and the *between-sector* level. Traffic variables should predict changes in workload within sectors over time. For example, increases in aircraft count or proximity within a sector should produce an increase in workload. The structural properties of the airspace, in interaction with the traffic within the sector, should produce differences in workload between sectors. Random coefficient models allow us to examine the way in which variables at these two levels interact with each other. A separate regression equation is estimated for each sector. The parameters from these within-sector models (i.e., the intercepts and slopes) are then examined to assess whether they vary across sectors. If there is significant variance in the intercepts or slopes, between-sector variables (e.g., sector volume) can be specified as predictors of these parameters. In this way we can test whether the structural properties of the airspace modify the effect of traffic variables, thereby allowing us to predict how changes in sector design will affect workload.

Method

Procedure and Sample

Workload data was collected from 18 enroute radar sectors in Brisbane Centre, Australia. Ratings were provided by licensed air traffic controllers, who sat at a console adjacent to the sector under observation. The raters were instructed to observe the traffic in the sector, and assess the workload that a typical controller would experience if he or she was controlling that sector. Ratings were provided at two-minute intervals over a 45 minute period. Sixteen civil air traffic controllers provided ratings, all of whom held a current endorsement for the sector that they were rating. A total of 2,156 workload ratings were obtained. Radar track data, flight plan data and controller interactions with the system were extracted from the Eurocat system for the period under observation.

Measures

Workload was assessed using the Air Traffic Workload Input Technique (ATWIT; Stein, 1985). Raters were asked to assess the workload experienced by the controller over the past two minutes (W_t) on a 10-point Likert scale (1 = “Low workload: could accomplish everything easily”; 10 = “Extreme workload: extremely difficult to accomplish everything, assistance would be needed”).
A range of workload metrics were extracted from the radar track data. The current analyses incorporate four measures:

- Number of aircraft under jurisdiction (N_t);
- The ratio of the number of aircraft on descent to the number of aircraft under jurisdiction (D_t);
- Horizontal proximity (H_t); and
- Sector volume (V).

Horizontal proximity was calculated as the inverse of the minimum horizontal separation between aircraft pairs within an altitude band (metric C9 from Chatterji & Sridhar, 2001).

Results

Descriptive Statistics

Table 1 shows the means, standard deviations and correlations among the variables. As can be seen in Table 1, workload was positively correlated with aircraft count, the descent ratio, horizontal proximity and volume.

Table 1: Means, standard deviations and correlations among variables

Variable	M	SD	W_t	N_t	D_t	H_t
W_t	2.40	1.09				
N_t	4.26	2.38	.41*			
D_t	0.26	0.27	.11*	-.17*		
H_t	0.02	0.06	.25*	.33*	-.06*	
V	439836	452324	.26*	.34*	-.03	-.02

Note: * p<.05

Multi-level Analyses

Analyses for this study were conducted using Hierarchical Linear Modeling (Bryk & Raudenbush, 1992). The dependent variable was the workload rating at a given point in time (W_t). Two levels of analysis were included in the model: the within-sector level (level 1), and the between-sector level (level 2). The independent variables at the within-sector level were aircraft count, descent ratio, and horizontal proximity. Sector volume was an independent variable at level 2, because it is a property of the sector as a whole. In order to obtain other measures of sector-level properties, we aggregated aircraft count, descent ratio, and horizontal proximity to the sector level. We did this by calculating the averages for each of these measures for each sector (*Av[N]*, *Av[D]*, *Av[H]*). These aggregated measures reflect the nature of the traffic that typically flows through the sector. For example, approach sectors typically have a high ratio of aircraft on descent, and the aircraft tend to be closely spaced. All variables were uncentered.

The procedure recommended by Hofmann, Griffin and Gavin (2000) for multi-level modeling was then followed. The first step of this procedure involves running a model with no predictor variables (an 'empty model') in order to estimate the percentage of variance in workload that resides at the two levels. This analysis revealed that 54% of the variance in workload was at the within-sector level, leaving 46% of the variance at the between-sector level. Thus, our results demonstrate that there is variability in workload both within and between sectors.

Step 2 involves running a model, in which the level 1 predictors are entered as random effects (an 'unconstrained model'). The only predictor with a reliable random effect was the number of aircraft under jurisdiction. This demonstrates that the effect of aircraft under jurisdiction on workload varies across sectors. The random effects for descent ratio and horizontal proximity, by contrast were not reliable. For this reason, we subsequently fixed these two variables. The level 1 predictors accounted for 27% of the within-sector variance and 54% of the between-sector variance.

Step 3 involves running a model, in which level 2 variables are entered as predictors of the intercepts of the level 1 equations (an 'intercepts as outcomes model'). Because our variables are uncentered, the intercepts of the level 1 equations represent the amount of workload that a controller has when there are no aircraft in the sector. From a practical perspective, differences in this parameter are not meaningful. For this reason, we fixed the intercepts and skipped step 3.

If there is reliable variance in the slopes of the level 1 equations, then predictors of the slopes are entered at Step 4. As noted above, there was reliable variance in the effect of aircraft count, so we attempted to account for this variance. We initially tried entering volume, average aircraft count, average descent ratio, and average horizontal proximity as predictors. However, the model would not run when volume and average aircraft count were included. The final model that we ran is below:

$$W_t = \pi_0 + \pi_1 N_t + \pi_2 D_t + \pi_3 H_t + e$$

$$\pi_0 = \beta_{0,0}$$

$$\pi_1 = \beta_{1,0} + \beta_{1,1} Av[D] + \beta_{1,2} Av[H] + r$$

$$\pi_2 = \beta_{2,0}$$

$$\pi_3 = \beta_{3,0}$$

The results of this final model are presented in Table 2. As can be seen in Table 2, aircraft count predicted workload at the within-sector level, but the descent ratio and horizontal proximity did not. However, the average descent ratio and the average horizontal proximity of aircraft in the sector did moderate the effect of aircraft count. The effect of aircraft under jurisdiction was stronger in sectors that typically have a higher descent ratio, and aircraft in closer proximity. Average descent ratio and proximity accounted for 38% of the variance in the effect of aircraft count across sectors.

Table 2. Final model

Fixed effect	Coefficient	*SE*
Intercept, π_0		
Intercept, $\beta_{0,0}$	1.624*	0.043
Aircraft count, π_1		
Intercept, $\beta_{1,0}$	-0.005	0.062
Mean descent ratio, $\beta_{1,2}$	0.698*	0.249
Mean horizontal proximity, $\beta_{1,3}$	$3.739^{\dagger}$	2.104
Descent ratio, π_2		
Intercept, $\beta_{2,0}$	0.081	0.070
Horizontal proximity, π_3		
Intercept, $\beta_{3,0}$	-0.237	0.317

Note: * $p<.05$; † $p<.10$

To illustrate the differences in effect of aircraft count, scatter plots for Keppel/Alma (Figure 3) and Daintree/Reef-Low (Figure 4) were examined. Superimposed over the scatter plot, we show the regression line for that sector. The regression line was calculated using the mean descent ratio and proximity values for the two sectors, and the coefficients in Table 2 ($\beta_{1,0}$, $\beta_{1,2}$, $\beta_{1,3}$). The scatter plots show that the relationship between aircraft count and workload is stronger in Daintree/Reef-Low than in Keppel/Alma, and that the HLM model captures this effect reasonably well.

Despite the fact that these two sectors are similar in size and average traffic load, the difference in effect of aircraft count on workload can be explained by the different traffic patterns in these sectors. In Daintree/Reef-Low, a larger proportion of the aircraft is on climb out of or on descent into an airport. This has 2 consequences; a) it is more difficult for controllers to maintain separation assurance for aircraft that are changing levels (Loft et al., 2007); and b) especially during times of higher traffic load, controllers will have to maintain flow sequence for the aircraft on descent into an airport. In Keppel/Alma, the aircraft are mostly flying level and the route structure provides separation. Southbound aircraft fly on a one-way route to the west of the sector, whereas the northbound aircraft fly on a one-way route further east. In addition, the jet and the propeller traffic are naturally separated by their preferred respective cruise flight levels. Most of the work for Keppel/Alma controllers is monitoring whether or not faster following aircraft are catching up to the aircraft in front of them. This example shows how different factors trigger workload in different sectors and that the average workload for an additional aircraft is less for a structured sector like Keppel/Alma than for a more complex sector like Daintree/Reef-Low.

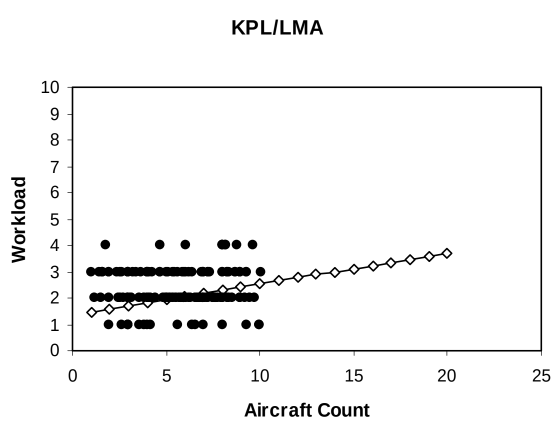

Figure 3. Effect of aircraft count in Keppel/Alma

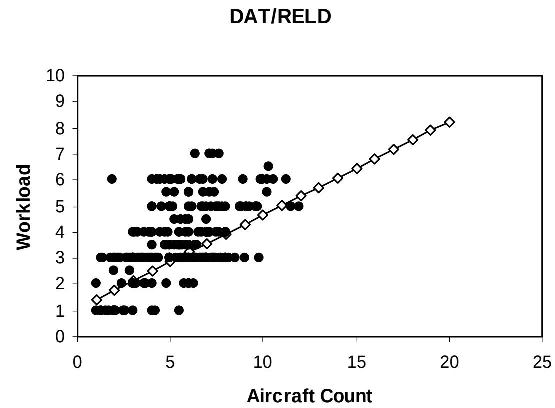

Figure 4. Effect of aircraft count in Daintree/Reef-Low

Discussion

The current results show that the effect of aircraft count varies across sectors, and that differences in the effect of aircraft count can be predicted from the nature of the traffic that typically flows through the sector. It is interesting that the descent ratio and proximity of aircraft did not predict workload at the within-sector level, yet they did moderate the effect of aircraft count at the between-sector level. Thus, it is not merely the case that aircraft that are on descent, and which are in close horizontal proximity, require greater monitoring by controllers. If this were true, it would produce effects for descent ratio and proximity at the *within-sector* level, rather than the *between-sector* level. The results show that after controlling for the descent ratio and proximity of aircraft at a given point in time, a single aircraft in Daintree/Reef-Low (for example) will still impose a greater workload than a single aircraft in Keppel/Alma.

There are several reasons why aircraft may impose higher workload on controllers in sectors where the majority of aircraft are typically on descent, and are typically closely spaced. We believe that the primary reason is because sectors that score highly on these variables are mostly approach sectors. Controllers in these sectors have additional work involved in assigning flight levels, providing separation assurance, and meeting flow requirements.

The current findings have implications for the development of workload metrics. The results confirm the belief that a generic 'one-size fits all' workload metric will not work. The parameters that predict workload vary across sectors, and we cannot assume that a workload model that has been fitted to data collected in one sector will be able to predict workload in another sector. This is a problem, because one of the reasons why Air Navigation Service Providers (ANSP) need workload metrics is because they want to be able to use them for sector redesign. If a group of sectors is redesigned by changing routes and boundaries, then it will change the emergent properties of the traffic. For example, changes to the boundaries or routes may change the proportion of aircraft that are typically on descent, or the proximity of those aircraft. Our results suggest that a workload model calibrated on the original sector may not generalize to the new sector.

Random coefficient modeling provides one way of addressing this problem. By partitioning the variance in workload into within and between-sector components, we were able to disentangle the effects of sector properties from the effects of the traffic within that sector. We were thus able to identify what features it was about the sector as a whole that were responsible for differences in the effects of within-sector variables. This model allows us to make predictions regarding the effects of changes in sector design. For example, if Centre Managers were to consider shifting the southern boundary of Keppel/Alma further south this would increase the average descent ratio and average horizontal proximity of aircraft, as there is another aerodrome to the south of the current sector boundary. Thus, the model would predict that effect of aircraft count would become stronger. If so, then the maximum number of aircraft that could be safely handled in that sector would decrease.

In summary, our findings show that it is possible to develop workload metrics that are sensitive to differences in sector properties. This approach to workload modeling identifies a required commitment on the part of ANSP to collecting workload data in the operational environment. There are clearly challenges in collecting this type of data. However, if ANSP want to be able to identify, assess, mitigate and manage the risk associated with sector re-design, then the development of predictive workload metrics must be undertaken.

References

Arad, B. A. (1964). The controller load and sector design. *Journal of Air Traffic Control (May)*, 12-31.

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing and illustrations. *Organizational Research Methods, 5(4)*, 362-387.

Boag, C., Neal, A., Loft, S., & Halford, G. (2006). An analysis of relational complexity in an air traffic control conflict detection task. *Ergonomics*, 49, 1508–1526.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. California: Sage Publications.

Chatterji, G. B., & Sridhar, B. (2001). Measures for air traffic controller workload prediction. *Proceedings from the 1st AIAA Aircraft Technology, Integration and Operations Forum*, Los Angeles, CA.

Davis, C. G., Danaher, J. W., & Fischl, M. A. (1963). *The influence of selected sector characteristics upon ARTCC controller activities*, Contract No. FAA/BRD-301, Arlington, Virginia: The Matrix Corporation.

Jolitz, G. D. (1965). *Evaluation of a mathematical model for use in computing control load at ATC facilities* (Report No. RD 65-69). Atlantic City, NJ: Federal Aviation Administration.

Hofmann, D. A., Griffin, M. A., & Gavin, M. (2000). The application of Hiearchical Linear Modeling to management research. In K. Klein & S. Kozlowski (Eds.), *Multilevel Theory, Research, and Methods in Organizations*. San Francisco: Jossey Bass

Kopardekar, P., & Magyarits, S. (2003). Measurement and Prediction of Dynamic Density. In *Proceedings of the 5th USA/Europe ATM 2003 R & D Seminar*, Budapest, Hungary.

Leiden, K. J., Kopardekar, P., and Green, S. (2003). Controller workload analysis methodology to predict increases in airspace capacity. AIAA's 3rd Annual Aviation Technology, Integration, and Operations (ATIO) Tech, Denver, Colorado.

Loft, S., Bolland, S., & Humphreys, M. (2007). Modeling the human air traffic controller, Part I: Expert-trainee differences in conflict detection. To appear in *Proceedings of the 14th International Symposium on Aviation Psychology*.

Loft, S., Sanderson, P. M., Neal, A. & Mooij, M. (in press). A critical review of modelling and predicting mental workload in en-route air traffic control. *Human Factors*.

Manning, C., Mills, S., Fox, C., & Pfleiderer, E. (2001). *Investigating the validity of performance and objective workload evaluation research (POWER).* (Report No. DOT/FAA/AM-01/10). Washington, DD: FAA Office of Aviation Medicine.

Schmidt, D. K. (1978). A queuing analysis of the air traffic controller's work load. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-8(6),* 492-498.