Delft University of Technology

Cross-sector transferability of metrics for air traffic controller workload

Abdul Rahman, Mariam; Borst, Clark; van Paassen, Rene; Mulder, Max

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Cross-sector transferability of metrics for air traffic controller workload

S.M.B. Abdul Rahman * , **C. Borst** ** , **M.M. van Paassen** **
and M. Mulder **

* *Mechanical Engineering Faculty, Universiti Teknologi MARA, 40450*
Shah Alam, Selangor Darul Ehsan, Malaysia
(e-mail: mariam4528@salam.uitm.edu.my)
** *Control and Simulation section, TU Delft, Delft, The Netherlands*
(e-mail: {c.borst, m.m.vanpaassen, m.mulder}@tudelft.nl)

Abstract: Air traffic controller workload is an important impediment to air transport growth. Several approaches exist that aim to better understand the causes for workload, and models have been derived to predict workload in new operational settings. These methods often relate workload to the difficulty, or complexity, that an average controller would have to safely manage all traffic in a sector with a particular traffic demand. In this paper, several of these complexity-based metrics for workload will be compared. Of special interest is whether the complexity measures *transfer* from one sector design to another. That is, does a metric that is well-tuned to predict workload for controllers working in one sector, also predict the workload for another group of controllers active in a different sector? Results from a human-in-the-loop experiment show that a solution space-based metric, which requires no tuning or weighing at all, has the highest correlations with subjectively reported workload, and also yields the best workload predictions across different controller groups and sectors. *Copyright © 2016 IFAC*

Keywords: Air traffic control, taskload, mental workload, supervisory control

1. INTRODUCTION

Safety, efficiency and orderly flow of air traffic are the three main Air Traffic Controller (ATCO) responsibilities in managing traffic. Current Air Traffic Control (ATC) practice primarily uses conventional technology (e.g., radar and radio telephony communication), with only little automation support for the operators involved, which renders the task of supervising air traffic heavily constrained by human performance limits (Costa, 1993). Without countermeasures, the rise in projected air traffic would inevitably result in a further increase in the workload of ATCOs, often cited as one of the main impediments to air transport growth (Janic, 1997, Hilburn, 2004, Koros et al., 2004).

The ability to understand what causes workload, and predict ATCO workload in future scenarios, is an important avenue of research. In this paper we use the term taskload to refer to the objective demands of a task, and workload to address the subjective demand as experienced by an operator (Stassen et al., 1990). Several approaches exist to determine ATC taskload, such as simply counting the number of aircraft that need to be managed simultaneously in a sector. Although this technique works quite satisfactorily, it does not include any knowledge regarding *how* these aircraft fly through the sector. Figure 1 illustrates that a situation where all aircraft fly parallel routes is very likely to be much easier for an operator to supervise and control than a situation where the *same* number of aircraft fly random routes.

More recent techniques relate task demand load to metrics of sector complexity (Laudeman et al., 1998, Sridhar et al., 1998, Chatterji and Sridhar, 2001, Kopardekar and Magyarits, 2002, Masalonis et al., 2003). An important example is the dynamic density (DD) metric, which includes aircraft dynamic behavior in the sector, by taking into account "*the collective effort of all factors or variables that contribute to sector-level ATC complexity or difficulty at any point of time*" (Kopardekar and Magyarits, 2002). The DD calculation is based on weights that are gathered from applying regression methods on samples of traffic data and comparing these to subjective workload ratings. The DD metric therefore includes both objective as well as subjective measurements and could be less suitable to predict the workload of different controllers working in another sector.

In the solution space (SSD)-based approach, taskload is related to the difficulty of the ATC control problem, where the "solution space" captures the geometrical and
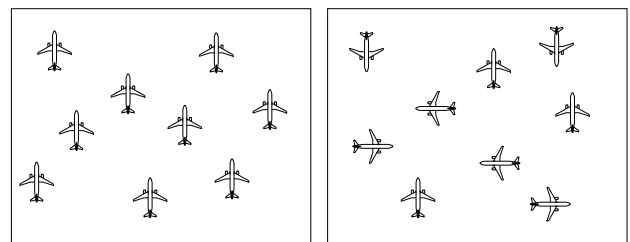

Fig. 1. Two traffic situations, with the same number of aircraft, one easy and the other difficult to control.

kinematic constraints that limit (and therefore, guide) ATCO control actions (Hermes et al., 2009, Mercado-Velasco et al., 2010, D'Engelbronner et al., 2015). Previous studies found high correlations between workload ratings and the *area* of the available SSD control space.

This paper discusses a comparison of several sector complexity measures regarding their ability to match the subjective workload ratings obtained in a human-in-the-loop experiment. We will evaluate the Static Density (SD), which equals the number of aircraft flying in a sector, the Dynamic Density (DD) as proposed by NASA, and a solution space-based (SSD) approach developed by TU Delft. We will focus in particular on the performance of these metrics in predicting workload ratings *across* different sectors and *across* different groups of operators, i.e., their ability to transfer between sectors and controllers.

2. EXPERIMENT

Our study relies on computing the correlation between ATCO workload ratings and a number of complexity metrics: SD, SSD and DD. A human-in-the-loop experiment was conducted in which eight participants, who all received an extensive ATC introductory course and has worked closely in the ATC domain, but none of them were operational ATCOs, managed the air traffic in two sectors (Abdul Rahman, 2014). While managing the air traffic, every minute the subject was requested to indicate the workload on a scale between 0 and 100, yielding a workload profile for each controller. After each run, based on the recorded aircraft parameters (their position, speed, and heading), the complexity metrics were computed, and for the DD metrics the weightings were determined through linear regression techniques. When all data were available, the correlation analysis was conducted.

2.1 Method

Independent variables   The experiment had two independent variables: (i) two different sector designs were used, Figure 2, and (ii) four different traffic sequences were simulated. The latter were varied to avoid scenario recognition during the course of the experiment.

The two sectors differed in the number of crossing points, combinations of the intercept angle of traffic routes, the clustering of crossing points, different entry and exit points, differences in sector shape and sector area. The four traffic patterns did not differ in the total number of aircraft simulated, but rather in their distribution in time.

In addition, we divided the eight participants in two groups of four subjects each, to allow us to study the effects of using the metrics across groups of participants.

Subject instructions   Subjects were instructed to guide all aircraft safely through the sector and have them exit the sector at their pre-defined exit point. All aircraft were of the same type, so had the same constraints in velocity and heading; altitude was fixed to one flight level.

Procedure   All subjects were briefed on the nature of the experiment, the goals to be achieved and the simulator used. Each participant completed two blocks of
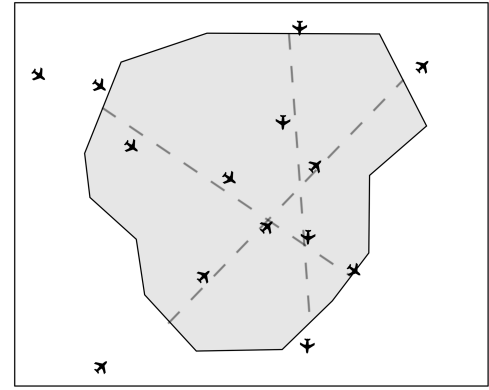
(a) Sector 1
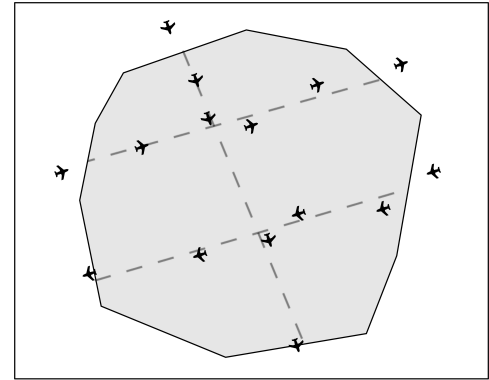
(b) Sector 2

Fig. 2. Sector design and traffic flows.

four scenarios that lasted 25 minutes each. Each block was preceded with a training scenario that lasted for ten minutes. Subjects were asked to indicate their workload using a scale that appeared on top of the plan view display. The workload rating, measured on a zero to 100 scale, was provided by the subject every 60 seconds during the experiment run. In order to correct for inter-subject differences, Z-scores of the subjective ratings were used in the subsequent data exploration. This correction was performed by calculating the Z-scores for every test subject.

The experiment was run at four times real-time, similar to what was done in previous research (Hermes et al., 2009, d'Engelbronner et al., 2010, Mercado Velasco et al., 2010). The rationale behind this was to create more variability in traffic situations (and thus workload) within relatively short experimental scenarios.

Dependent measures   Many variables have been collected, but here only the workload ratings, and the complexity metrics introduced above will be briefly discussed; see (Abdul Rahman, 2014) for details. Note that to rule out any 'fade in' and 'fade out' effects, the first 3 minutes and the last 2 minutes of each 25 minutes run were excluded, Figure 3.

The SD metric is equal to the total number of aircraft (N_{ac}) that fly through the sector, computed every minute. The SSD metric used was the mean area of the SSD of all aircraft in the sector, computed every minute (Hermes et al., 2009). Two DD metrics were computed: the $NASA_1$
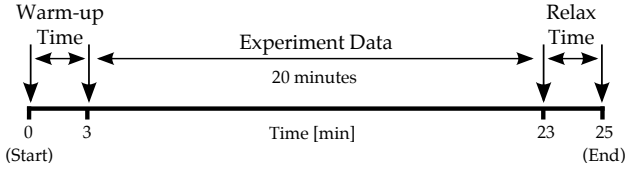
Fig. 3. Period where data is gathered in the experiment.

and NASA$_2$ metrics introduced in (Chatterji and Sridhar, 2001) and (Laudeman et al., 1998, Sridhar et al., 1998), respectively:

$$DD_{NASA_1} = \sum W_i DV_i, \quad \text{and} \quad DD_{NASA_2} = \sum W_i DV_i + TD,$$

with W_i the weights, DV_i the so-called dynamic variables used in the two metrics, and TD the traffic density (N_{ac} divided by the sector area).

The main difference between the NASA$_1$ and NASA$_2$ metrics lies in the choice of the DVs: for NASA$_1$ the DVs included the number of aircraft, the horizontal proximity, etc., whereas for NASA$_2$ the DVs included the number of heading changes, speed changes, etc. See (Chatterji and Sridhar, 2001) and (Laudeman et al., 1998, Sridhar et al., 1998) for more details.

The DD metrics were used in two different ways: (1) all weights W_i were set to '1', yielding the 'unweighted' DD metrics, and (2) the weights were calculated through a linear regression that fitted the DD to the subjective workload ratings, resulting in the 'weighted' DD metrics.

Every minute, we obtain the workload rating, and values for the SD, SSD area and the two unweighted and two weighted DD metrics. The workload ratings were first (on a per-subject basis) transformed to Z-scores. Then the correlation coefficients were computed between the Z-scores time series and the individual SD, SSD and DD metrics, using Kendall's tau (test statistic R).

2.2 Hypotheses

Based on our previous work, we hypothesized that, overall, the SSD metric would result in the highest correlations with ATCO subjective workload ratings. The weighted DD ratings, however, could surpass the SSD correlation quality as here the linear regression could optimize the weightings W_i for the measured DV's for the sector and group of participants being analyzed. However, when using exactly these weightings then for *another* sector, and/or *another* group of participants, we expect that the correlations would be lower again, revealing that the SSD-based metric is a less scenario- and subject-dependent metric.

3. RESULTS

3.1 Effects of the four traffic sequences

Figure 4 shows the mean number of aircraft (the SD) in the two sectors, as a function of simulation time, for the four traffic sequences. The figure shows that, on average, the traffic density in the sectors was independent of the traffic sequence; traffic density is higher in Sector 2. Statistical tests showed that none of the dependent measures were significantly affected by the traffic sequence.
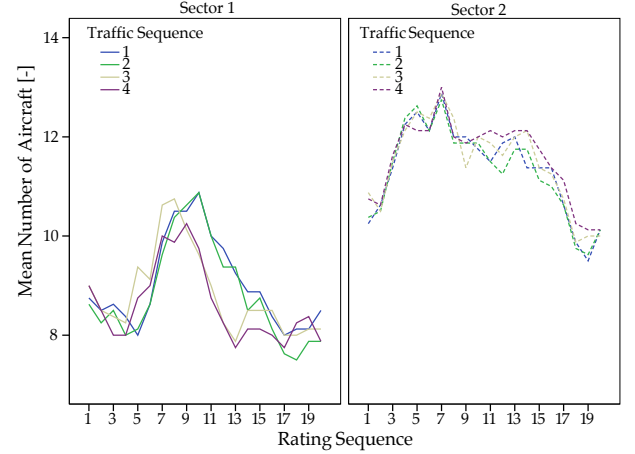

Fig. 4. Traffic density in the two sectors.

In the following, we can therefore focus on examining the differences between the two sectors, taking all traffic sequences together.

3.2 Effects of the two sectors

To analyze the transferability of our metrics from one sector to another, it is important that both sectors represent different levels of complexity. Figure 5 shows the total number of the three possible ATCO commands (speed, heading, speed+heading) and the number of times an aircraft was selected, for both sectors. Overall, more commands were given in Sector 1, which on average had a smaller number of aircraft (see Figure 6), a significant effect ($p=0.012$), but which had a more complex design.
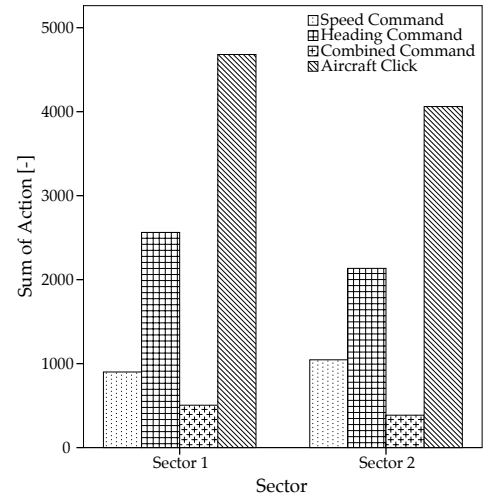
Fig. 5. Sum of the number of commands.

Indeed, the average workload ratings for Sector 2 were lower, Figure 7(a), a significant effect ($p=0.0125$). Figure 7(b), Figure 8(a) and Figure 8(b) show the averages of the SSD area metric and both NASA DD metrics, respectively. Clearly, these metrics were lower for Sector 2 (all significant at $p=0.012$). From this analysis we conclude that Sector 2 was indeed significantly easier to control than Sector 1. It illustrates that our intention to create a sector with *more* aircraft, but which was *easier* to control because of a lower complex sector design, was indeed successful.

(a) total number a/c (b) average number a/c

Fig. 6. Number of aircraft (total, average).

(a) Z-scores (b) SSD area

Fig. 7. Averages of the workload ratings, SSD area metric.
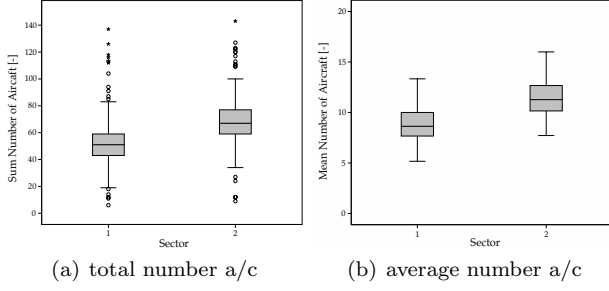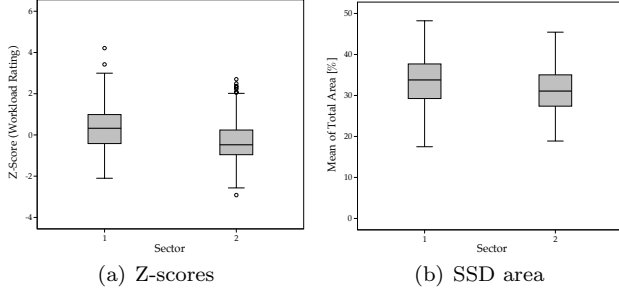
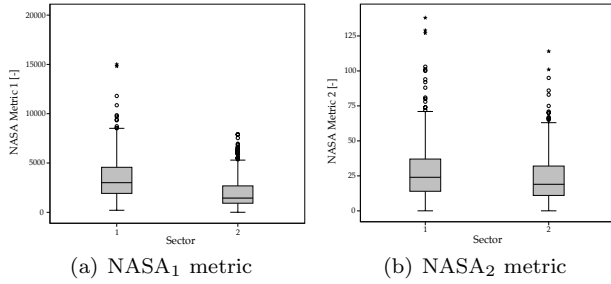This will be a good test for the transferability of the metrics, discussed next.

(a) NASA$_1$ metric (b) NASA$_2$ metric

Fig. 8. Averages of the NASA$_1$, NASA$_2$ metrics.

3.3 Unweighted correlation analysis

Sector-based analysis   Results of a correlation analysis between the number of aircraft, the unweighted NASA DD and the SSD metrics with respect to the ATCO workload ratings is summarized in Table 1. The SSD metric yields the highest correlations, for both sectors, of all metrics. There is a striking difference between the performance of NASA$_1$ and NASA$_2$ for Sector 2. NASA$_1$ appears to be more sensitive to a change in sector layout and traffic structure than the other metrics (see also Figure 9). However, whereas NASA$_2$ and traffic density seems unaffected by sector changes, the SSD metric results in a better correlation with workload for Sector 1. This suggests that the SSD would be better in predicting workload when traffic streams are "less organized" with crossing points close together.

Group-based analysis   A similar result was observed when looking at different groups of controllers, where the SSD showed the highest correlation with the workload ratings and less sensitivity to a change in controller group (see Table 2). Again, NASA$_1$ has a poor correlation with

Table 1. Correlations between workload ratings and complexity metrics, grouped by sector.

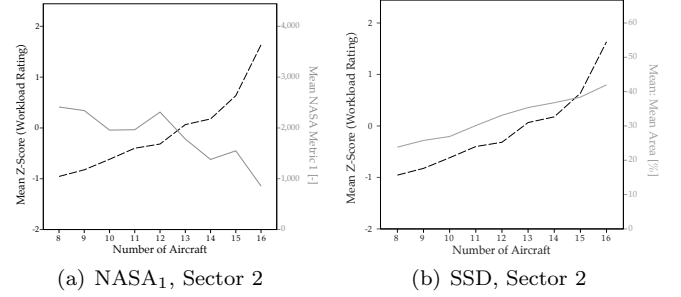|		$NASA_1$	$NASA_2$	SSD	N_{sum}	N_{mean}
Sector 1	R	0.170	0.256	**0.337**	0.215	0.297
	p	<0.001	<0.001	**<0.001**	<0.001	<0.001
Sector 2	R	-0.015	0.256	**0.290**	0.215	0.276
	p	0.564	<0.001	**<0.001**	<0.001	<0.001

(a) NASA$_1$, Sector 2 (b) SSD, Sector 2

Fig. 9. Correlation of NASA$_1$ and SSD for Sector 2.

workload and is also most affected by controller group. Interestingly, the remaining metrics all have a relatively lower correlation coefficient for controller Group 2, who were working with Sector 2. There is, however, no clear explanation for this result.

Table 2. Correlations between workload ratings and complexity metrics per controller group.

	$NASA_1$			
	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	0.178	0.010	-0.069	0.090
p	<0.001	0.785	0.066	0.016

	N_{sum}			
	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	0.233	0.214	0.265	0.162
p	<0.001	<0.001	<0.001	<0.001

	SSD			
	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	**0.362**	**0.335**	**0.341**	**0.232**
p	**<0.001**	**<0.001**	**<0.001**	**<0.001**

	$NASA_2$			
	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	0.269	0.261	0.306	0.204
p	<0.001	<0.001	<0.001	<0.001

	N_{mean}			
	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	0.335	0.284	0.336	0.209
p	<0.001	<0.001	<0.001	<0.001

3.4 Weighted correlation analysis

Sector-based analysis   In general, the NASA DD metrics should improve their correlations with subjective workload when the weights (per dynamic variable) are determined through regression analysis. In Table 3 it can indeed be seen that especially the NASA$_1$ metric improved considerably compared to the unweighted version (see Table 1) and has even a higher correlation than the SSD metric, for Sector 1. Surprisingly, NASA$_2$ slightly deteriorates in its performance as a workload predictor for Sector 1 as compared to the unweighted case.

Group-based analysis   The weighted NASA$_1$ metric surpasses the SSD metric in its correlation with subjective workload for Group 1 and Sector 1 (see Table 4). NASA$_2$ is more similar to the SSD correlations over all groups, except for Group 2 working with Sector 1. Here, NASA$_2$ has a slightly higher correlation coefficient compared to the SSD metric. Overall, it is clear that the two DD variants improve on their ability to predict workload when the complexity factors are weighted, where the weights are based upon regression analysis with subjective workload.

Table 3. Correlations between workload ratings and weighted DD metrics, grouped by sector.

	$NASA_1$		$NASA_2$	
	Sector 1	Sector 2	Sector 1	Sector 2
R	**0.375**	0.266	0.190*	**0.296**
p	**<0.001**	<0.001	<0.001	**<0.001**

*correlation at a lower level than unweighted NASA DD metric

Table 4. Correlations between workload ratings and weighted DD, per controller group.

	$NASA_1$				$NASA_2$			
	Sector 1		Sector 2		Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
R	**0.429**	0.229	0.217	0.111	0.358	**0.352**	**0.348**	**0.312**
p	**<0.001**	<0.001	<0.001	0.003	<0.001	**<0.001**	**<0.001**	**<0.001**

3.5 Transferability analysis

Cross-sector analysis   To analyze the sensitivity of the weighted DD metrics in terms of correlation to workload, a cross analysis between the two sectors was carried out. That is, the $NASA_1$ weights gathered for Sector 1 were applied to Sector 2 and vice versa. The same was done for the $NASA_2$ metric and the results are summarized in Table 5. From this table it can be seen that the correlation slightly deteriorates from the values listed in Table 1, except for $NASA_2$ in Sector 1. There is no logical explanation for this apparent increase in correlation for this case. To assess the SSD's sensitivity to sector changes, Figure 10 shows that the distribution of data points for all participants in Sector 1 and Sector 2 are almost identical, implying a relatively low sensitivity to sector changes.

Table 5. Correlations between workload ratings and cross-sector weighted DD metric.

	$NASA_1$		$NASA_2$	
	Sector 1	Sector 2	Sector 1	Sector 2
R	0.230	0.231	0.317*	0.245
p	<0.001	<0.001	<0.001	<0.001

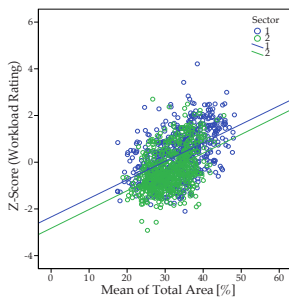*correlation at a higher level than weighted NASA DD metric

Fig. 10. Workload ratings versus the SSD metric.

Cross-group analysis   Similar to the cross-sector sensitivity analysis, a cross-group analysis was performed by applying the weights gathered for Group 1 to Group 2 and vice versa. The results as summarized in Table 6 reveal similar findings as the cross-sector results, namely that in certain conditions (i.e., group and sector) the correlations improve and in others the correlations worsen. This random behavior clearly indicates the DD's sensitivity to a change in controller group. Note that the randomness in correlation coefficients could be caused by outliers in

the data distribution. In general, outliers can significantly affect the goodness of fit. Figure 11 shows little difference between the workload-SSD relationship per controller group within a specific sector, whereas the data points are more distributed for Sector 1.

Table 6. Correlations between workload ratings and cross-grouped weighted DD metric.

	$NASA_1$				$NASA_2$			
	Sector 1		Sector 2		Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
R	**0.366**	0.264*	0.114	0.136*	0.358	**0.343**	0.129	0.241
p	**<0.001**	<0.001	0.002	<0.001	<0.001	**<0.001**	<0.001	<0.001

*correlation at a higher level than weighted NASA DD metric

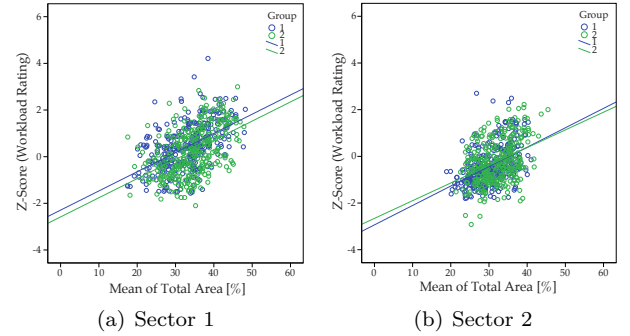(a) Sector 1   (b) Sector 2

Fig. 11. Workload ratings versus SSD area per group and sector.

4. DISCUSSION

This paper compared the solution space-based SSD-metric with established metrics such as the number of aircraft (SD) and NASA's dynamic density DD (Laudeman et al., 1998, Sridhar et al., 1998, Chatterji and Sridhar, 2001). Multiple scenarios with two different sectors and with varying traffic sequences were presented to subjects.

First, an analysis with regards to the subjects' overall behavior and workload ratings was conducted, to observe whether both sectors represent different complexity levels, a necessary condition for our cross-sector transferability investigation on workload metrics. The two sectors indeed represented different complexity levels, with the sector with the least aircraft to be in fact more complicated to manage. All sector characteristics, such as its area, route design and location of route intercept and sector entry/exit points contributed to the effort needed to control it. This is consistent with the current practice to define the maximum number of aircraft on a *per sector* basis.

Initial correlation analyses were conducted to compare the SSD-based metric and the unweighted DD metrics towards the workload rating. The analysis aimed at having a neutral comparison, that is, without the effects of post-processing procedures such as weighting the DD coefficents using linear regression. As hypothesized, the SSD metric had the highest correlations with workload, relative to the unweighted DD metrics and also the number of aircraft. This is found for both sectors, and for both groups of controllers.

Then, the DD metrics were 'tuned', through establishing an optimal set of weighting coefficients to yield the best

relation between the DD and the workload ratings. Different sets of DV weightings were used to tune the DD metric for the two sectors and each individual group group of controllers. It was found that the differences in controller's strategies lead to different weightings, showing that the DD metric is affected by the controllers for which the metric is tuned. Relative to the unweighted DD metrics, overall the correlations with workload improved, and some weighted metrics even got higher correlation values than the SSD metric.

However, when *transferring* a particular DD model, optimized for one sector and one group of controllers, to a different sector or different group of controllers, the correlations were again lower. This clearly indicates that, in contrast to the SSD-metric which is independent of sector and controllers, the DD metric is sensitive and therefore less suitable to predict ATCO workload in different sectors and with different controllers as compared to the sector and controllers for which the metric was obtained.

Note that the original DD metrics, however, were constructed based on three-dimensional airspace, with traffic samples from 36 high and low sectors. The linear regression analysis in this paper computed the DD metrics based on a two-dimensional, rather simplified airspace, and also using a low number of participants. Therefore, there could have been the possibility that our DD 'models' were being overfit; minor fluctuations in our data could have deteriorated the metrics' performance. Nevertheless, the DD metrics should not be too sensitive to a specific sample size and should perform well on any sector design or group of controllers.

5. CONCLUSIONS

This paper showed that the solution space-based complexity metric (SSD) is a more reliable and objective sector complexity measure, as compared to the static and dynamic density metrics. It managed to show the same high level of correlation with ATCO subjective workload under various air traffic sector designs and for different groups of controllers. The SSD metric can be used in real-time without any post-processing procedures, potentially allowing for a real-time prediction of ATCO workload. It should be noted, however, that these results were gathered with regards to specific assumptions and experiment settings. To prove that the constraint-based method using the SSD metric is the most suited metric in measuring sector complexity construct in a real operational setting, a more extensive research regarding its performance and robustness should be done.

ACKNOWLEDGEMENTS

REFERENCES

Abdul Rahman, S.M.B. (2014). *Solution Space-Based Approach to Assess Sector Complexity in Air Traffic Control*. Ph.D. thesis, Aerospace Engineering, TU Delft.

Chatterji, G.B. and Sridhar, B. (2001). Measures for Air Traffic Controller Workload Prediction. In *1st AIAA Aircraft Technology, Integration, and Operations Forum*, 1. AIAA, Los Angeles, CA.

Costa, G. (1993). Evaluation of Workload in Air Traffic Controllers. *Ergonomics*, 36(9), 1111–1120.

d'Engelbronner, J., Mulder, M., van Paassen, M.M., de Stigter, S., and Huisman, H. (2010). The Use of the Dynamic Solution Space to Assess Air Traffic Controller Workload. In *AIAA Guidance, Navigation, and Control Conference*, 21. AIAA, Toronto, CA. AIAA-2010-7542.

D'Engelbronner, J.G., Borst, C., Ellerbroek, J., Van Paassen, M.M., and Mulder, M. (2015). Solution Space Based Analysis of Dynamic Air Traffic Controller Workload. *Journal of Aircraft*, 52(4), 1146–1160.

Hermes, P., Mulder, M., Van Paassen, M.M., Boering, J.H.L., and Huisman, H. (2009). Solution Space-Based Analysis of Difficulty of Aircraft Merging Tasks. *Journal of Aircraft*, 46(6), 1995–2015.

Hilburn, B.G. (2004). Cognitive Complexity in Air Traffic Control - A Literature Review. Technical Report EEC Note 04/04, EUROCONTROL, Brétigny-sur-Orge, France.

Janic, M. (1997). A Model of Air Traffic Control Sector Capacity Based on Air Traffic Controller Workload. *Transportation Planning and Technology*, 20, 311–335.

Kopardekar, P. and Magyarits, S. (2002). Dynamic Density: Measuring and Predicting Sector Complexity. In *Proc. of the 21st IEEE/AIAA Digital Avionics System Conference*, 1–9. IEEE, Pascataway, NJ.

Koros, A., Della Rocco, P., Panjwani, G., Ingurgio, V., and D'Arcy, J.F. (2004). Complexity in Air Traffic Control Towers: A Field Study. Technical report, NTIS, Springfield, VA.

Laudeman, I.V., Shelden, S.G., Branstron, R., and Brasil, C.L. (1998). Dynamic Density: An Air Traffic Management Metric. Technical Report NASA-TM-1998-112226.

Masalonis, A.J., Callaham, M.B., and Wanke, C.R. (2003). Dynamic Density and Complexity Metrics for Realtime Traffic Flow Management. Technical report, Center for Advanced Aviation System Development, MITRE.

Mercado-Velasco, G.A., Mulder, M., and Van Paassen, M.M. (2010). Air Traffic Controller Decision-Making Support using the Solution Space Diagram. In *Eleventh IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Valenciennes, France, August 31 - September 3*, 227–232.

Mercado Velasco, G.A., Mulder, M., and van Paassen, M.M. (2010). Analysis of Air Traffic Controller Workload Reduction Based on the Solution Space for the Merging Task. In *AIAA Guidance, Navigation, and Control Conference*, 18. AIAA, Toronto, CA. AIAA-2010-7541.

Sridhar, B., Sheth, K.S., and Grabbe, S. (1998). Airspace Complexity and its Application in Air Traffic Management. In *2nd USA/Europe Air Traffic Management R&D Seminar*, 1–9. Orlando, FL.

Stassen, H.G., Johannsen, G., and Moray, N. (1990). Internal Representation, Internal Model, Human Performance and Mental Workload. *Automatica*, 26(4), 811–820.