ATC Complexity as Workload and Safety Driver

Jelena Djokic, Bernd Lorenz
EUROCONTROL CRDS
Budapest, Hungary
jelena.djokic@eurocontrol.int,
bernd.lorenz@eurocontrol.int

Hartmut Fricke
Institute of Logistics and Aviation
Technical University of Dresden
Dresden, Germany
fricke@ifl.tu-dresden.de

***Abstract*— This paper describes an investigation into ATC complexity as a contributory factor in changes of safety level. ATC complexity, together with equipment interface and procedural demands comprise the task demands on the controller; subsequent controller activities are mediated by performance shaping factors to create workload. In order to establish a link between ATC complexity, a controller's subjective workload and safety, complexity factors are identified and subsequently related to both workload and safety indicators. The studied data comes from a real-time simulation using controller-pilot data-link communication (CPDLC) technology, recently completed at EUROCONTROL CRDS in Budapest.**

Keywords - ATC complexity; task demands; controller's activity; workload; safety

I. INTRODUCTION

The EUROCONTROL Statistics and Forecast Service (STATFOR) predicts that in 2025 the number of commercial flights in Europe will be between 1.7 and 2.1 times the traffic in 2005 [1]. This is an average growth of 2.7%-3.7% per year. The most pressing problem facing the European Air Traffic Management (ATM), therefore, will be to provide sufficient capacity to meet this increased air traffic demand, while at the same time the safety level of air travel has to be maintained or even improved. Airspace capacity that lags behind air traffic demand inevitably leads to flight delays, which in turn means an economic loss to airlines.

In the current air traffic control (ATC) environment the key limiting factor to increasing sector capacity is the workload of the air traffic controller. Therefore proposed solutions for increasing airspace capacity aim at reducing controller workload - which includes: the delegation of separation tasks from ground to the aircraft (e.g. the free-flight concept [2]), a re-sectorisation of the airspace, and the introduction of new controller support tools in order to reduce the amount of work, or at least the difficulty of the controller tasks. As the work of air traffic controllers is foremost cognitive in nature a considerable amount of research has been undertaken to understand the complex task demands that drive the workload of a controller (see [3] for a recent review). The term "workload" denotes a subjective quality reflecting the individual controller's perception of the task demand imposed on him/her by the current air traffic situation. Thus, many studies implicitly assume that controller workload varies as a function of both directly measurable air traffic factors (number of aircraft in the sector, speed variability, proximity of aircraft, etc.) and controller's activity mediated by factors such as the controller's abilities, age, fatigue, level of experience, etc. [4].

As ATC is a safety-critical working environment any changes implying an adverse impact on controller workload have a direct bearing on flight safety (Fig.1.adopted after [4]).

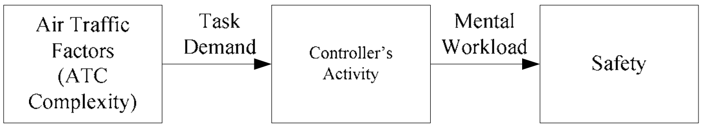

Figure 1. Simplified scheme of the relationship between ATC complexity and safety

The present paper examines the relationship between task demands as defined by a set of ATC complexity factors, controller's actions, subjective workload, and safety. For a safety criterion, a metric was used that was recently developed as part of the EUROCONTROL INTEGRA project [5]. This metric referred to is Propensity and is defined as a proxy for the likelihood of a safety significant event occurring during normal operations. Thus, the present study attempts to predict subjective workload and Propensity as criteria on a moment-to-moment basis using a linear combination of ATC complexity factors and controller's activity measures as predictors.

The paper is organized as follows: In the following section we will give a brief overview of research on ATC complexity and the derivation of task demand metrics from which a selection for the purpose of the present study was made. A more detailed description of controller's activity and workload measures used in the study follows. Then, the INTEGRA safety metric is explained. Next is described the real-time simulation experiment which provided the data base for the calculation/collection of the predictor and criterion metrics. What follows are a description of the approach for statistical analysis and the presentation of the results. Finally, these results will be discussed and conclusions drawn.

This work is funded by the EUROCONTROL CEATS Research, Development and Simulation Centre (CRDS), Budapest, Hungary.

ISBN: 978-0-615-20720-9

II. ATC COMPLEXITY

A straightforward determinant of controller workload is simply the number of aircraft for which the controller is responsible in a specified time and sector. This measure is referred to as the sector load. Predicting sector load and avoiding sector overload is the basic tool upon which current traffic flow management is built. However, the level of difficulty experienced by the controllers depends on additional factors beyond the number of aircraft present in a sector [6]. To be able to capture more accurately ATC complexity, it is necessary to take into consideration also flight characteristics of each individual aircraft as well as interactions between pairs of aircraft. Important flight characteristics of aircraft relate to instantaneous changes of the state of the aircraft, e.g. changes in altitude, heading or speed. Interactions between aircraft are considered not only in terms of potential conflicts but also include the pattern of how aircraft converge and the degree of what in [7] has been referred to as the disorder among aircraft, i.e. the variability in headings and speeds of aircraft. Despite the fact that ATC complexity has been the subject of a significant number of studies (see [3] for a recent review), and many complexity factors have been proposed, up to now a comprehensive and generally accepted set of measures has not been defined yet.

For the purpose of the present study, a list of complexity factors was selected that has been consistently found to be important and for which detailed calculation formula have been reported. The factors were partially elicited from work described in [8 - 12].

The selected overall set of 24 complexity factors is presented in Table I. It is out of the scope of this paper to describe all 24 factors in detail. For a more thorough review of the listed factors readers are referred to the indicated source literature.

III. CONTROLLER ACTIVITY – LINK BETWEEN TASK DEMANDS AND CONTROLLER'S WORKLOAD

Even though task demand factors can capture one aspect of the ATC situation, it should be kept in mind that ATC is a dynamic environment and that controllers actively interact with the traffic, and therefore have an important influence on ATC complexity and hence the level of safety.

Several researchers agree that workload is a result of such a complex interaction between the task demand and the way the controller actively manages the situation (e.g. [3], [4], [13], [14]). Moreover controllers, by performing certain activities, regulate the evolution of the task demands with the aim of keeping workload at an acceptable level.

Nevertheless, not all controller tasks are observable. As defined by [15], there are four controller tasks while managing the ATC situation: monitoring, evaluating, planning and implementing the formulated plan. Furthermore, out of these four tasks only one is observable, and that is the implementation process. It means that by taking only objectively measurable (sub)tasks into consideration, it is possible to capture only one aspect of comprehensive controller activity involved. However, as this aspect of the controller's activity is directly connected with changes made by the controller on the ATC situation, we considered it sufficient for our study.

Thus, in the current study, controller's input (data entries) and radio communication were used as the representatives of performed controller's activities (the study is based only on the Executive controller data entries, and not Planning controller, and therefore no phone communication is not considered here).

TABLE I. COMPLEXITY FACTORS

	Complexity Factors	Used in
1	number of aircraft	[8], [10], [12]
2	number of climbing aircraft	[8], [10], [12]
3	number of descending aircraft	[8], [10], [12]
4	number of aircraft with heading change greater than 15˚	[10], [9]
5	number of aircraft with the speed change greater than 10 knots	[10], [9]
6	number of aircraft with lateral distance between 0-25nm and vertical separation less than 2000ft above 29000ft	[10], [9]
7	horizontal proximity measure 1 (C5)	[8], [10], [12]
8	vertical proximity measure 1 (C6)	[8], [10], [12]
9	horizontal proximity measure 2 (C7)	[8], [10]
10	vertical proximity measure 2 (C8)	[8], [10], [12]
11	horizontal proximity measure 3 (C9)	[8], [10]
12	vertical proximity measure 3 (C10)	[8], [10]
13	time-to-go to conflict measure (C13)	[8], [10]
14	variance of ground speed (C14)	[8], [10], [12]
15	ratio of standard deviation of speed to average speed (C15)	[8], [10], [12]
16	density indicator (mean)	[7], [11], [12]
17	variability in headings (track_disorder)	[7], [11], [12]
18	variability in speed (speed_disorder)	[7], [11], [12]
19	divergence between pairs of aircraft	[7], [11], [12]
20	convergence between pairs of aircraft	[7], [11], [12]
21-22	sensitivity indicator (a/c converging-mean; a/c diverging-mean)	[7], [11], [12]
23-24	insensitivity indicator (a/c converging-mean; a/c diverging-mean)	[7], [11], [12]

IV. THE INTEGRA CONCEPT OF SAFETY METRICS

Within the INTEGRA concept [16] the term Propensity expresses the likelihood of a safety significant event occurring during the operation of the ATM system. It is defined through the probability function of an aircraft about its calculated position. Therefore, the interactions between aircraft are presented through the interactions between these probability functions. Propensity is calculated for each pair of aircraft that are within defined cut-off criteria for both vertical and horizontal separation.

If the distance between two aircraft is decreasing, the

 ISBN: 978-0-615-20720-9

interaction between their probability functions will be higher, and therefore, the value of propensity metric is increasing. Additionally, the INTEGRA authors introduced into propensity calculation a so-called safety weighting function for the pair of aircraft, determined by the distance between the two aircraft. The purpose of the safety weighting function is to describe the safety significance of proximity between aircraft. This weighting function also takes the use of advanced tools, the density of air traffic, higher than normal information processing loads and severe weather into account.

The propensity defined for an aircraft pair takes values between 0 and 1: the propensity tends to 1 when there is a major reduction in safety margins and tends to 0 when safety margins are assured. More detailed information on the propensity metric can be found in the [16].

V. The Real-Time Simulation Experiment

1) Simulation

In order to obtain relevant values, data were recorded during a two-week LINK2000+ Small Scale Real Time Simulation 2 experiment (LINK 2000+ SSRTS2). The aim of this simulation was to develop and validate new principles of task delegation between the planning and executive controller with the aim to best accommodate the Controller-Pilot Data-Link Communication (CPDLC) capability in an en-route environment [17, 18]. The simulation involved three different sectors of the Central European Air Traffic Services (CEATS) airspace. The data used for the present study are data obtained for the two busiest sectors simulated.

2) ATC Complexity measures

The flight plans and flown trajectories were used as input data for the complexity factors calculations. Customized software was developed to calculate these values for each 2 minute time steps.

3) Controller activity measures

All inputs made by the executive controller recorded during the simulation were extracted. These inputs refer to assignments of vertical rate, exit flight levels/planned entry levels, cleared flight levels, headings, speed instructions, and direct clearances. These were summed across each 2-minute time step and across input, resulting in only one measure named Actions_SUM. Furthermore, cumulative duration of radio calls (= frequency occupancy time per 2-minute time step) was calculated as well as the average duration of single calls. Altogether, we used three measures of the controller's activity – Actions_SUM, Frequency Occupancy Time and Average Radio Duration obtained for every 2-minute time steps.

4) Workload measures

For the same time steps, controllers were providing workload ratings. To collect workload measures during the simulation the Instantaneous Self Assessment (ISA) technique as operator-subjective metric was applied, where the air traffic controller gives subjective ratings of workload. This tool was developed by the UK NATS and offers 5 points rating scale. On every time step controller can opt a level of workload ranging from very low to very high.

5) Safety measure

As noted before, the propensity metric was used as the safety criterion. The propensity value is calculated using software developed within the frame of the EUROCONTROL INTEGRA project. The values are obtained for each pair of aircraft within a sector unit during the 2-minute time steps. Bearing in mind that there was more than one pair of aircraft in the sector, it was necessary to extract one value per time step. In the study we opted for the maximum value.

Descriptive statistics of the extracted measures are given in table II.

TABLE II. Descriptive Statistics of dependent measures

Measures	Min.	Max.	Mean	Std. Deviation
Actions_SUM (count)	0.000	13.000	3.320	2.372
Frequency Occupancy Time (s)	0.000	50.400	23.991	8.779
Average Radio Duration (s)	0.000	7.200	3.245	0.674
ISA (rating 1 to 5)	1.000	5.000	2.870	0.736
propensity_max (value 0 to 1)	0.010	0.976	0.599	0.147

6) Participants and data extraction procedure

The LINK2000+ SSRTS2 experiment involved a total of 18 controllers out of which 6 controllers worked on the two sectors considered here. The data used for the statistical analysis were derived from these 6 participants only. Each controller completed an overall of 8 exercises of 1 hour and 20 minutes, from which 1- hour recordings were extracted for analysis. Scores were derived for every 2 minutes, resulting in 30 measurements per exercise. These data were obtained for each indicator (ATC complexity measures, workload measures and safety measures). The overall dataset comprised 6 (controllers) x 8 (exercises) x 30 (time segments) = 1440 measurements for each indicator. Prior inspection of the data set revealed that during the whole simulation (all 8 exercises) one of the six controllers always rated workload as 'fair', hence there were no variations in workload measure. The data of this participant was discarded from the analysis, with 1200 measurements remaining. In 58 time segments (4.8%) data was missing. Therefore, subsequently reported results are based on measurements obtained in 1142 time segments.

VI. Statistical Evaluation and Results

A. Principal Component Analysis

In a first analysis step, a Principal Component Analysis (PCA) on all 24 complexity metrics was computed in order

ISBN: 978-0-615-20720-9

to derive a reduced number of uncorrelated predictor variables for the subsequent computation of regression models.
Principal components having an eigenvalue > 1 were extracted and subsequently rotated using the VARIMAX method. This analysis resulted in the extraction of 8 principal components that accounted for 67.26 % of the total variance in the metrics. The table III. displays these components sorted by the sizes of their eigenvalues and along with the percentage of variance they account for. By inspection of the pattern of loadings given in the 8 component x 24 metrics matrix, the following components meanings could be derived. Note, that the loading of a given metric on a given component is equivalent to the correlation between that metric and that component. Therefore, the metric with the highest loading by and large guides the interpretation of the component.

TABLE III. RESULTS OF THE PRINCIPAL COMPONENT ANALYSIS

Components	Eigenvalue	% of Variance	Cum. % of Variance
Comp.1	4.935	20.563	20.563
Comp.2	3.442	14.343	34.906
Comp.3	1.775	7.395	42.301
Comp.4	1.450	6.042	48.343
Comp.5	1.344	5.602	53.944
Comp.6	1.129	4.706	58.650
Comp.7	1.036	4.316	62.966
Comp.8	1.030	4.291	67.257

Comp.1 – ground speed variance and *divergence/convergence*: strongly related to the variance of the ground speed (0.884) and the ratio of the standard deviation to the mean ground speed (0.845). Also, the strong correlation with divergence and convergence factors (0.787 and 0.785 respectively) was recognised, which is in compliance with speed significance, as divergence/ convergence factors actually measure how fast aircraft are moving toward/from each other.

Comp. 2 – aircraft count: this component has the strongest correlation with the number of the aircraft in the sector (0.816)

Comp. 3 – horizontal proximity: this component can be considered as addition to the previous one, as it shows high correlation with the horizontal distance between aircraft taking into consideration the aircraft count - horizontal proximity measure (C5) : 0.894 and density_mean: 0.815 .
Together these two components (*Comp. 2* and *Comp. 3*) can be representatives of so-called *sector density*.

Comp. 4 – aircraft vertical transitioning: highly correlated to the number of descending aircraft (0.785) as well as speed change related to this vertical evolution (0.732)

Comp.5 – conflict sensitivity: this component is loaded highly by both sensitivity indicators ($Sd_+(i)$:0.772 and $Sd_-(i)$: 0.751). Sensitivity is related to the gradient of the relative distance between aircraft. This indicator measures the change in terms of relative distance in response to changes in speed and heading of the involved aircraft. If sensitivity is high only small changes in heading and speed imply a high impact on relative distance. This is the case, e.g. when two aircraft are heading towards each other. The sensitivity indicators are designed to set a weight on potential conflicts that are difficult to solve (see 12]. Note that a situation with high sensitivity is easier to resolve for the controller than one with a low sensitivity [7].

Comp.6 – insensitivity: This component is strongly related to the insensitivity indicators both for convergence and divergence of the aircraft (*insen_c*: 0.723 and *insen_d*: 0.686). It is not simply an analogue with the opposite direction to the previous component. High insensitivity is given for a pair of aircraft when the degree of convergence is high while sensitivity for convergence is low.

Comp.7 – vertical separation: high correlation with the measure of the vertical separation of aircraft in close horizontal proximity (C10) defines this component (0.849)

Comp.8 – horizontal separation: analogously to the previous component, this component is defined based on the correlation with the measure of horizontal separation of the aircraft in close vertical proximity C9 (0.908).
The PCA yielded 8 component scores for each two-minute interval which were used as predictors in the subsequent multiple regression analyses.

B. Multiple Regression Analyses

Two sets of multiple regression models were computed. The first set was performed to assess the effectiveness in predicting ISA workload ratings on the basis of ATC complexity and controller activity metrics. The second set of multiple regression analysis assesses the effectiveness of predicting propensity using ATC complexity, controller activity and ISA workload ratings as predictors.

1) ISA regression models

Instead of using the stepwise linear regression involving all predictors, we first compared two alternative multiple regression equations to fit the data. For the first multiple regression equation all 8 component scores were forced into the model regardless of their single significance. In the second equation the 3 activity metrics entered the equation. This was done in order to assess the contribution of ATC complexity components in relation to the controller activity metrics. Table IV contains the global statistics of these two equations. It can be seen that the first equation containing only complexity components yielded a multiple R of 0.36

ISBN: 978-0-615-20720-9

corresponding to R^2 of 0.13 or in other words corresponding to 13% of variance of the ISA workload ratings. Adding the controller activity measures in the second equation contributed to a significant increase in the multiple R although it increased the percentage of variance explained by only 3% to a total of 16%. Therefore, it can be concluded that both sources of information, ATC complexity and controller activity, have a unique contribution to the prediction of ISA workload ratings.

TABLE IV. COMPARISON OF ALTERNATIVE MULTIPLE REGRESSION MODELS FOR PREDICTION OF ISA

Regression equation containing	mult. R	R^2	R^2 change	F change	df	Sig. F change
complexity components	0.36	0.13	0.13	21.20	8 , 1131	0.000
complexity components and controller's activity measures	0.40	0.16	0.03	13.77	3 , 1128	0.000

A final multiple regression model was computed using traditional stepwise linear regression approach in order to identify those predictors that are responsible for the significant contribution to workload prediction. This model we refer to as the optimised model as all insignificant variables have been removed. The parameter statistics of this model is given in Table V. The model consists of 8 parameters (Comp.1 – Comp.6, Frequency Occupancy Time and Average Radio Duration)

TABLE V. PARAMETER STATISTICS OF THE OPTIMISED MODEL FOR THE PREDICTION OF ISA WORKLOAD RATINGS

	B	Std. Error	Beta	t	Sig.
Comp.1	0.052	0.020	0.071	2.596	0.010
Comp.2	0.102	0.021	0.139	4.921	0.000
Comp.3	0.101	0.020	0.137	5.009	0.000
Comp.4	0.092	0.020	0.124	4.561	0.000
Comp.5	-0.147	0.021	-0.200	-7.166	0.000
Comp.6	0.074	0.020	0.100	3.676	0.000
Frequency Occupancy Time	0.012	0.003	0.142	4.627	0.000
Average Radio Duration	-0.182	0.033	-0.164	-5.488	0.000

The stepwise regression analysis revealed that the first 6 out of 8 complexity components remained in the prediction model. The components that showed the strongest correlation with ISA ratings are Comp. 3 and Comp. 5 which consider horizontal proximity and the conflict sensitivity. The higher horizontal proximity, i.e. the closer the aircraft in the horizontal plane, the higher was controller workload. When sensitivity of the conflict increased, the workload ratings of the controller decreased, which is consistent with [7].

Frequency Occupancy Time and Average Radio Duration representing the communication load also remained in the model. When Frequency Occupancy Time, i.e. overall frequency occupancy time is increasing, the workload rating is also higher. On the other hand, the increment of the average duration shows the decrease in controller's workload.

2) Propensity regression models

Analogously to the ISA regression models, we compared alternative multiple regression equations in order to evaluate how the three sources of indicators (ATC complexity, controller activity and ISA workload) contributed to the prediction of propensity. Four propensity regression equations were considered: the first equation contains only the scores of the ATC complexity components, two "intermediate" equations in addition contain the three activity measures or the ISA rating, respectively. The fourth equation contains all input variables. The global statistics of these equations are shown in table VI. The equation that contains only ATC complexity components yields a multiple R of 0.53, i.e. accounts for 28% of variance of the propensity metric. When comparing the two "intermediate" equations, one can see that adding the 3 activity measures to equation 1 improved the prediction only by 1% (see R^2 change in equation 2 in table VI), while adding ISA improved the prediction by 2%. However, there is no gain in predictive power when the measures of controller activity are added to equation 3 as indicated by the statistics obtained for equation 4, the full model.

TABLE VI. COMPARISON OF ALTERNATIVE MULTIPLE REGRESSION MODELS FOR PREDICTION OF PROPENSITY

Regression equation containing	mult. R	R^2	R^2 change	F change	df	Sig. F change
1.complexity components	0.53	0.28	0.28	55.99	8, 1133	0.000
2. complexity components and controller's activity measures	0.54	0.29	0.01	3.63	3, 1130	0.013
3. complexity components and ISA ratings	0.55	0.30	0.02	26.05	1, 1130	0.000
4. complexity components, ISA and controller's activity measures	0.55	0.30	0.00	1.83	1, 1127	0.139

Finally, stepwise regression analysis was performed for the identification of an optimized model. The parameter statistics this model are shown in table VII. The model consists of 5 parameters (Comp.2 – Comp. 6 and ISA ratings).

TABLE VII. PARAMETER STATISTICS OF THE OPTIMISED MODEL FOR THE PREDICTION OF PROPENSITY METRIC

	B	Std. Error	Beta	t	Sig.
Comp. 2	0.027	0.004	0.184	7.268	0.000

 ISBN: 978-0-615-20720-9

Comp. 3	0.041	0.004	0.280	11.104	0.000
Comp. 4	0.014	0.004	0.095	3.795	0.000
Comp. 5	-0.049	0.004	-0.332	-12.924	0.000
ISA ratings	0.028	0.005	0.138	5.192	0.000

The Comp.1, which stands for ground speed variance and divergence/convergence of the pair of aircraft, did not remain in the model. Even though through Principal Components Analysis it resulted in highest eigenvalue and accounted for 20.563% of the total variance in the metric, it did not contribute to the prediction of propensity metric. The components that showed the strongest correlation with propensity metric are Comp. 3 and Comp. 5, i.e. horizontal proximity and the conflict sensitivity. Also, of great significance are correlations of propensity metric with number of aircraft in the sector (Comp.2) and vertical movements of the aircraft in the sector (Comp.4). As it could be anticipated through the previous comparison of alternative regression models, the activity measures did not remain in the model. Only ISA ratings contributed to the prediction of propensity.

VII. Discussion and Conclusions

The focus of the paper was the investigation of the relationship between ATC complexity, controller's activity measures, subjective workload and safety measures. Based on the previous work in the field an initial set of 24 complexity factors was defined. In order to reduce this set, a Principal Component Analysis (PCA) was performed, which resulted in 8 components. [12] also performed a PCA using a set of 27 complexity indicators as input variables. There is a big overlap between their and our set of input variables (see complexity factors that is used also in [12] in table I). However, their data was extracted in one-minute time steps from real traffic recorded in a total of 103 sectors across one day of traffic. Their PCA revealed 6 components (using the same extraction criterion of eigenvalue > 1 as in the present paper) that accounted for 76% of the total variance. Aircraft count had the highest loading on the first component accounting for 46.7% of the variance in their PCA which corresponds to the second component in the PCA of the present study. Comparing the loadings of the complexity indicators on the remaining five components suggests that the first component of our PCA corresponds to a mix of their second, fourth and fifth component. Finally, their sixth component is more or less equivalent to our seventh component. Their third component in [12] was correlated with a metric representing the degree of incoming sector flows, which was not considered here. Therefore, a good agreement between the PCA results obtained in our simulation study and their real-traffic study can be concluded.

The scores for the eight components were calculated and further entered in different multiple regression models in order to reveal their correlation with ISA workload measures, controller activity indicators and INTEGRA propensity as a safety measure.

First of all, it was found that subjective controller workload as measured by the ISA ratings depends on additional factors rather than only on aircraft count. This is in agreement with a couple of other studies (e.g. [7], [9], [12]).

The results suggested that subjective workload hinges on other aspects of the ATC complexity as well as on the communication load of the controller. Both the total frequency occupancy time and average radio duration significantly correlate with ISA workload ratings. Moreover, as it was hypothesised in [19], the present study suggested that the average time for an individual communication is negatively related to workload. In other words, the amount of time that a controller spends on a single communication should decline as the situation gets busier.

Furthermore, the results of the propensity regression model provided some insight into the construct validity of this metric in the sense that we first can pinpoint what aspects of the sector situation influence the degree of propensity, and second, in what way propensity is related to subjective workload and measures of controller activity. Four complexity components were found to be correlated with propensity: aircraft count, horizontal proximity, vertical transitioning, and conflict sensitivity. It is assumed that it is due to the calculations of the propensity metrics which relies on the calculated position of aircraft rather than their speed, the latter being reflected in component 1 of the PCA.

It is still an issue for further validation to demonstrate that propensity is a valid safety metric. This should be taken into account in future work, which should consider also other the actual occurrence of safety critical events into account.

References:

[1] EUROCONTROL, "EUROCONTROL long-term forecast: IFR flight movements 2006-2025 (DAP/DIA/STATFOR Doc216)", EUROCONTROL Statistics and Forecast Service (STATFOR), 2006.

[2] RTCA, "Report of the RTCA board of directors' select committee on free flight", Washington DC, 1995.

[3] B. Hilburn, "Cognitive complexity in air traffic control - a literature review (EEC Note No. 04/04)", EUROCONTROL, 2004.

[4] S. Loft, P. Sanderson, A. Neal and M. Mooij, "Modeling and predicting mental workload in en route air traffic control: critical review and broader implications", in Human Factors: The Journal of the Human Factors and Ergonomics Society, 2007, vol. 49, pp. 376-399.

[5] EUROCONTROL, "INTEGRA metrics & methodologies - detailed specification of safety metrics; inputs-processing-outputs", EUROCONTROL, 2000.

[6] B. Sridhar, K.S. Sheth and S. Grabbe, "Airspace complexity and its application in air traffic management", In Proceedings of the 2nd USA/Europe Air Traffic Management 1998 R&D Seminar, Orlando, Florida, 1998.

[7] D. Delahaye and S. Puechmorel, "Air traffic complexity: towards intrinsic metrics", In Proceedings of the 3rd USA/Europe Air Traffic Management 2000 R&D Seminar, Napoli, Italy, 2000.

[8] G.B. Chatterji and B. Sridhar, "Measures for air traffic controller workload prediction", In Proceedings of the First AIAA Aircraft

ISBN: 978-0-615-20720-9

Technology, Integration, and Operations Forum, Los Angeles, CA, 2001.

[9] I.V. Laudeman, S.G. Shelden, R. Branstrom and C.L. Brasil, “Dynamic density: an air traffic management metric (NASA-TM-1998-112226), San Jose State University Foundation, San Jose, CA, 1998.

[10] P. Kopardekar and S. Magyarits, “Measurement and prediction of dynamic density”, In Proceedings of the 5th USA/Europe Air Traffic Management 2003 R&D Seminar, Budapest, Hungary, 2003.

[11] F. Chatton, “Etudes de nouvelles métriques de complexité de la circulation aérienne”, Master’s thesis, Ecole Nationale de l’Aviation Civile (ENAC), 2001.

[12] D. Gianazza and K. Guittet, “Evaluation of air traffic complexity metrics using neural networks and sector status”, In Proceedings of the 2nd International Conference on Research in Air Transportation (ICRAT), Belgrade, Serbia, 2006.

[13] Majumdar, , W. Ochieng, G. McAuley, J. M. Lenzi, and C. Lepadatu, “The factors affecting airspace capacity in Europe: a cross sectional time-series analysis using simulated controller workload data.”, in Journal of Navigation, 2004, vol. 57, pp. 385–405.

[14] W. S. Pawlak, C. R. Brinton, K. Crouch and K. M. Lancaster, ”A framework for the evaluation of air traffic control complexity”, In Proceedings of the AIAAGuidance, Navigation and Control Conference, San Diego, CA, 1996.

[15] J.M. Histon and R.J. Hansman, “The impact of structure on cognitive complexity in air traffic control (Report No. ICAT-2002-4)”, MIT International Center for Air Transportation, Department of Aeronautics & Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA, 2002.

[16] EUROCONTROL, “INTEGRA metrics & methodologies - safety metrics; technical definitions”, EUROCONTROL, 2000.

[17] EUROCONTROL, “LINK2000+ small scale real-time simulation No.2 (LINK2000+ SSRTS2), final report (CRDS Note No.26)”, EUROCONTROL, 2007.

[18] R. Schuen-Medwed, B. Lorenz and S. Oze, “Empowerment of the planning controller in CPDLC environment”, In Proceedings of the Human Factors and Ergonomics Society, Europe Chapter Annual Meeting., Braunschweig, Germany, 2007

[19] C. Manning, S. Mills, C. Fox, E. Pfleiderer and H. Mogilka, “The relationship between air traffic control communication events and measures of controller taskload and workload”, In Proceedings of the 4th USA/Europe Air Traffic Management 2001 R&D Seminar, Santa Fe, 2001

ISBN: 978-0-615-20720-9