A Systematic Literature Review on Complexity–Safety Correlation Models Towards Safety-Conscious Arrival Management

Stanley Förster and Hartmut Fricke
Chair of Air Transport Technology and Logistics
Technische Universität Dresden
Dresden, Germany
stanley.foerster@tu-dresden.de

***Abstract*—The link between airspace and air traffic complexity and controller workload has been of high interest since decades, since a negative correlation is postulated between workload and safety. However, exploratory review indicated that this correlation is hardly quantified in the literature. In this study, we present a literature review following the systematic *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* workflow to ensure a good coverage and deep understanding of the relevant publications on this topic. We compile a search term out of three groups to set the scope of the review, which is the investigation of complexity and safety correlation in the air traffic control domain. After screening and reviewing over 1,900 publications obtained by querying four big research databases (IEEE Xplore, Scopus, Web of Science, and Semantic Scholar), we reviewed 72 papers in detail, six of which have been identified to attempt quantifying complexity – safety correlations. From these, we could learn that collision probability is not a suitable safety metric to correlate with complexity indicators in a predictive context, since corresponding work shows only little correlation. Rather, metrics closer to all-day operations of a controller, like loss of separation, are much more suitable.**

***Keywords*—air traffic control, air traffic management, complexity, workload, task load, safety**

I. Introduction

Despite the hit by the global COVID-19 crises, air traffic industry continuous to grow. Accommodating increasing traffic volumes requires both an optimized utilization of the available ATM capacity, and its extension for critical components such as the approach sectors of already busy airports. While new ground infrastructure such as additional runways or related high speed exits add capacity only locally and with very long response times, more effective measures consist of refined operations and tailored procedures. Some of these measures, like re-categorization of wake turbulence categories (now called wake turbulence groups, WTG by ICAO [1]), up to concepts including dynamic pair-wise separation, will, on the blind side, also increase heterogeneity of the traffic mix, tremendously. Innovative propulsion technologies for aircraft and new vehicle concepts may certainly support this trend. Consequently, the task load of air traffic controllers (ATCos) is supposed to increase as well, potentially at an extent that may compromise the safety of operations due to amplifying effects on the resulting workload. This is especially true in already very dense air spaces and at busy traffic times, respectively. Since safety is paramount to the societal acceptance of air traffic operations, it must be at least kept at given standards at all times. Therefore, introducing enhanced controller (decision) support tools mitigating traffic complexity and so assumingly task load is crucial to keep ATCo workload on an acceptable level.

Following this reasoning, we aim at developing a support tool that extends the capabilities of an Arrival Manager (AMAN) by generating additional suggestions to ATCos on exactly when to issue relevant clearances to the cockpit crew to reach improved throughput of dense airspace segments down to the runway. The goal is to integrate today's busy approach airspace segment, often representing a bottleneck to traffic flows, further into an entire flight covering safety oriented optimization process. Since control procedures are based on upstream safety assessment of route design, procedures and boundary conditions, in every-day operation, only operational performance indicators are being recorded and safety is considered to be granted using these indicators as proxies. This results in rather conservative and as such capacity-sensitive rules on aircraft separation minima, that do not allow leveraging the full system potential, especially not under all ambient and operational circumstances.

By integrating additional, real-time measurable performance metrics using a well tailored support tool extending today's AMAN concepts, the system will allow for continuous safety analyses on a tactical level. This way, we postulate, additional potential in airspace and runway capacities can be leveraged. A positive correlation is finally postulated between hazard potential and so-called *traffic complexity* as entropy, i.e. as a measure of traffic disorder and thus implicitly a measure of ATCo workload.

Before presenting recent scientific work related to correlation models, the following section gives a general overview of the current state of the art to air traffic complexity metrics and related safety research and motivates the need for an accordingly tailored systematic literature review. In section III we outline the methodology and workflow used to identify relevant publications. Afterwards, the actual process of querying, filtering, and evaluating publications is presented. Section IV gives an overview of our findings and highlights selected publications rated most relevant to the subject matter. The paper concludes with a discussion of the results and an outlook to anticipated next steps in this area.

II. State of the Art

The formal translation process from traffic complexity into workload follows a widely accepted concept depicted in figure 1. According to [2], *complexity* is an objective measure that allows quantifying the difficulty of the controller task, why it is also referred to as *task load*. How this translates into *workload*, which is a subjective measure of how an ATCo perceives this difficulty, is influenced by several factors, like the quality and capabilities of the controller working position (CWP) equipment, the behavioral profile of the present human, as well as individual strategies for handling those tasks. In turn, the level of workload indirectly influences complexity via controller behavior and the actions taken in front of the perceived workload, again (e.g. allow or deny additional aircraft to enter a sector).

The translation from complexity to workload will certainly be influenced by the availability of enhanced support tools and an increasing level of automation, since these affect the "equipment" part but also the strategies, how ATCos work. However, at the time being, the level of automation did not increase beyond the generation of advice and the human operator still stays responsible for checking and implementing the suggestions provided by the system [3]. Therefore, we decided to not restrict the selection of publications by requiring specific investigation of automation issues, yet. As can be seen from the results presented below, this decision proved reasonable, since automation aspects are hardly covered in previous work.

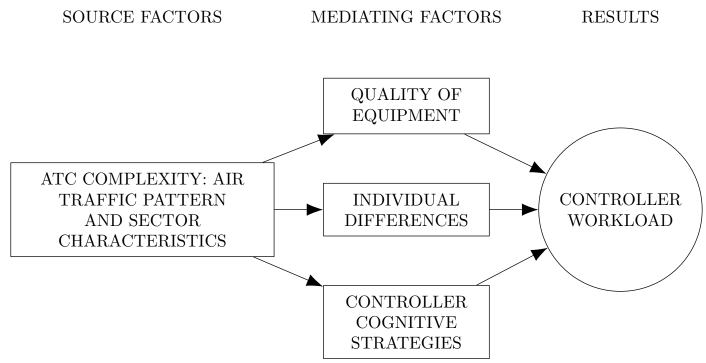

Figure 1. Transition from complexity to controller workload [2].

Investigating the state of the art of complexity-safety correlation models revealed that there is a lot of work available in which the correlation between objective indicators of complexity and subjective indicators of workload is investigated, as will be shown in detail in section IV. In most of this work, the correlation between complexity or workload and safety is assumed implicitly, not explicitly formulated. However, starting in 2012, the authors elaborated on how a correlation between complexity and safety can be formalized, c.f. figure 2 [4]. Back then, an investigation based on 19 complexity indicators from literature and collision probability did not unveil a significant statistical correlation. In fact, the authors concluded that the influence of workload is diffuse and that the most explainable safety metric is given by the collision probability.

Since then, only little progress could be found tackling the problem of quantifying the correlation between complexity and safety. To confirm, we take a step back to reiterate on this question and thoroughly investigate the current research landscape by systematically searching, reviewing, and evaluating publications that investigate the correlation between complexity and safety in the context of air traffic management.

III. Methodology

For the systematic literature review, we followed the ideas of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) workflow [5], as depicted in figure 3.

To begin with, a query has been constructed to search publication databases for relevant literature. The query is composed of three major keyword groups: First, the domain is defined (*air traffic control*). The second group comprises the independent variables of interest. Since the terms *complexity*, *task load*, and *workload* are used

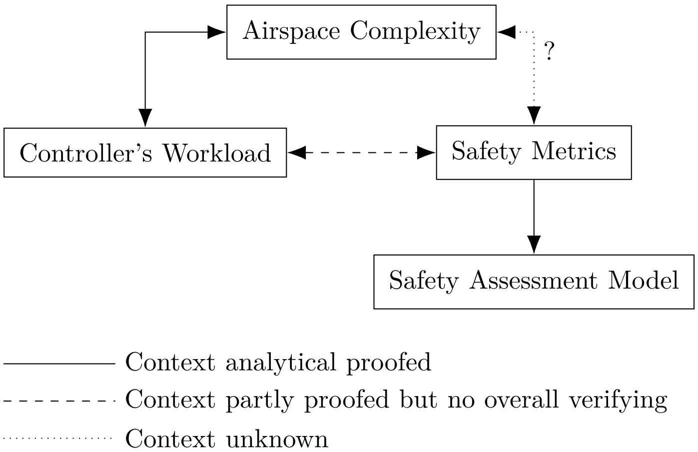

Figure 2. Relationship between complexity and safety metrics. Adapted from [4].

inconsistently throughout the literature, all three have been considered here. The third group states the dependent variables, i.e. safety related indicators and metrics. All three groups are connected via boolean `AND` operators, resulting in the following comprehensive search term: `( atc OR "air traffic control" OR atco OR "air traffic controller" OR atm OR "air traffic management" ) AND ( complexity OR workload OR "work load" OR taskload OR "task load" ) AND ( safety OR risk )`

For performing the query of publications, the following four databases have been selected: IEEE Xplore (330 results), Scopus (871 results), Web of Science (459 results), and Semantic Scholar (250 results). To include most recent publications, the search has been run multiple times, the latest update was performed on April 10, 2024.

After the final search, 1,910 publications were returned from the databases in total. Furthermore, 4 publications have been added that were not returned by any of the databases, based on external hints. A search for duplicates based on publication title and Digital Object Identifier (DOI) removed 592 records. The remaining records have been screened for eligibility. As criterion to include a publication for further investigation, we demand that the title or abstract has to give the impression that the work investigates or at least considers a correlation between workload or complexity, and safety. Furthermore, the work should address the air traffic domain. Even though, this has already been part of the search query, for some reason, the list of results included a lot of publications from other domains, especially health care, road traffic, and genetics. Using objective criteria like type of publication (e.g. only considering journal articles) or availability of a DOI was not feasible, since this would have excluded relevant work published at high-rated aviation conferences, like Air Traffic Management Research and Development Seminar (ATM Seminar), International Conference on Research in Air Transportation (ICRAT), and SESAR Innovation Days (SID). During this step, 1,130 records got removed from the results, leaving 192 publications for full paper screening.

For this, not all publications could be retrieved, due to access restrictions, unavailability, or language barriers. Among the results there was a single publication that got retracted. In summary, a set of 170 papers could be obtained for further processing. Eligibility for forwarding a paper to the final in-depth review step was decided based on the question, if it presents a quantitative correlation analysis or proposes a metric. Work that makes plain use of existing models for optimization or comparison of scenarios, concepts, tools, or algorithms, has been excluded. Additionally, the scope is narrowed to focus on air traffic control, i.e. work that focuses on cockpit side has been excluded. Finally, a set of 72 publications remained that were reviewed in detail as will be reported in the next section.

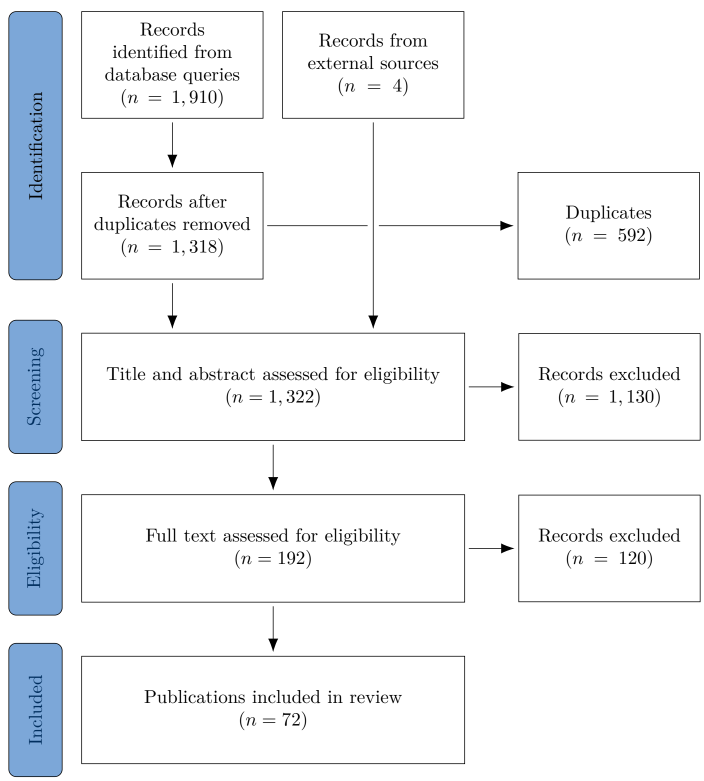

Figure 3. PRISMA workflow diagram outlining the selection process for publications considered for detailed review.

IV. Results

A. General Overview

During the in-depth review, our focus was to aggregate results and identify work that explicitly investigate indicator–safety correlations and quantify these by reporting coefficients of determination.

Throughout the final set of selected publications, the predominant way of considering safety aspects is to generalize that safety is negatively correlated to controller workload, i.e. increased levels of workload compromise safety, but without supporting this with appropriate studies, nor providing evidence for this assumption [6–37]. Another way is to related safety directly to complexity in general [12, 38–49], or to specific indicators like number of aircraft within a sector [50], again without supporting those assumptions.

Most studies – after mentioning these assumptions in their introduction – focus on investigating the correlation between selected traffic- or airspace-specific indicators and workload obtained in various ways. Widely used methods for the latter are human in the loop (HITL) studies, during which ATCos provide subjective measures of perceived workload via tools like Instantaneous Self Assessment (ISA), NASA Task Load indeX (NASA-TLX), Air Traffic Workload Input Technique (ATWIT), and Verbal Online Subjective Opinion (VOSO) [7, 11, 17, 21, 27, 30, 51, 52]. In some studies, instead of manual inputs, various psycho-/physiological and biochemical parameters, like electroencephalography (EEG), heart rate variability (HRV), Dual Frequency Head Maps (DFHM), eye movements, and reaction time, are measured during the simulation [7, 8, 13, 16, 21, 28, 29, 51–54]. Another way of obtaining subjective workload or complexity ratings to use for correlation analysis is through expert rating. Here, experienced ATCo get to rate scenarios based on recordings and replays without being situated in the actual role of an acting controller themselves [33, 46, 55]. Furthermore, Capacity analysis (CAPAN), and Directorate of Operation Research and Analysis Task (DORATASK) are two scores that focus more on a systematic analysis of controller tasks [26, 56].

Another group of publication comprises those in which authors propose metrics compiled from comprehensive sets of indicators that are supposed to reflect measures of complexity or workload, but without performing a correlation analysis to test whether those metrics reflect perceived workload of controllers [6, 15, 19, 20, 32, 35, 38–41, 45, 47, 48, 57–62].

Finally, some work relates complexity or workload indicators to capacity indicators, like number of aircraft passing an airspace, delay, or inter-arrival time [26, 63]

The following subsection gives detailed outlines of the publications that actually focus on safety indicators and -metrics.

B. Complexity and Safety Correlation Models

There exist several publications that focus on correlating indicators reflecting properties of the airspace or air traffic to safety-related indicators based on investigation of historic or simulation data. Safety indicators selected for analysis mostly comprise the number or frequency of incident like loss of separation, or near midair collision (NMAC) [64–66]. Often, the focus of a publication is on the qualitative cause analysis [67, 68]. An example of quantitative analysis is given in [65]. After the introduction of point merge operations at Istanbul Airport, authors compare two terminal maneuvering area (TMA) designs based on complexity as calculated using Eurocontrol's network strategic tool (NEST) tool, the number of aircraft within the airspace, as well as the number of conflicts (loss of separation as by NEST's definition). However, there is no correlation analysis between complexity and safety being performed.

In the following, we give an in-depth description of the most relevant papers, which actually relate complexity or workload indicators, and safety. Those have been selected based on the hard fact that they present coefficients of determination as proof of actual model development and correlation analysis. The order of the papers presented does not imply any preference.

1) Number of NMACs and Traffic Load: Gifford et al. develop a model to predict the number of NMACs from a linear combination of three metrics: Traffic Load, Complexity, and Staffing. The latter two are represented by linear functions of traffic load again [69]. Finally, the following model for the risk of an NMAC is presented:

$$RISK = \beta_0' + \beta_1' N + \beta_2' u + \beta_3 v + \epsilon \quad (1)$$

with

$$\beta_0' = \beta_0 + \beta_2\phi_0 + \beta_3(\gamma_0 + \phi_0\gamma_2)$$
$$\beta_1' = \beta_1 + \beta_2\phi_1 + \beta_3(\gamma_1 + \phi_1\gamma_2)$$
$$\beta_2' = \beta_2 + \beta_3\gamma_2$$

Here, Greek letters represent coefficients from the different sub-models of each metric. u, and v represent residuals of the complexity, and staffing metric, respectively, and N the number of aircraft within the sector. To validate the model, authors perform a data analysis of historic data from major US metropolitan regions, specifically 23 terminal control areas (TCAs) and 120 terminal radar service areass (TRSAs). For each type, the model is fitted separately, resulting in pseudo R^2 values of 0.61 for TCAs and 0.81 for TRSAs.

2) Probability of Operational Errors: In [70], authors correlate the number of aircraft within a sector to the probability of an operational error to occur. The model is based on the idea that each controller is subjected to a certain probability of making a control mistake. Hence, with each aircraft more within a sector, the chance of making a mistake increases proportionally. Data analyses lead to a probability of an operational error (per million control center-hours), depending on the number of aircraft n of ($R^2 = 0.989$)

$$P(\mathrm{OE}|n) = 2.6279 \cdot n^{2.479}. \tag{2}$$

3) Propensity: With the work presented in [71–73], authors aim at developing a linear model to correlate up to 24 complexity indicators, and ATCo activity measures with subjective workload ratings, as well as *Propensity*, a safety metric developed in the Eurocontrol INTEGRA project [74]. Controller activity measures comprise total and average duration of radio communication related tasks, as well as the number of inputs into the system. Workload is obtained during a HITL simulation study on a 5-level ISA scale. After performing a Principal Component Analysis (PCA) to reduce number of uncorrelated predictors, authors analyze the capabilities of different feature sets to predict ISA workload ratings and propensity scores, respectively. They find that the former has the highest correlation with a set of combined complexity, and controller activity indicators ($R^2 = 0.16$). The coefficients of a linear regression model predicting ISA scores are provided in [73]. Propensity correlates the most with a combination of complexity indicators and ISA ratings ($R^2 = 0.30$). The coefficients of the complexity-propensity model are provided in [71, 73]. Due to the large number of up to 32 coefficients, the models are not reflected here.

4) Collision Probability: In [4], authors search for a correlation between 19 complexity indicators, extracted from literature, and collision probability as safety metric [75]. Ahead of the data analysis, authors define 4 scenarios that differ in traffic load and mix (in terms of departure, and arrival) characteristics. Additionally, each scenario is split into 10 samples, each spanning a one-hour time frame. For each of the resulting 40 samples, a multiple regression analysis is performed separately, and selected results are highlighted. In summary, authors find that there is no significant correlation between complexity and collision probability. However, they conclude that the correlation is stronger for the TMA sector investigated ($0.37 < R^2 < 0.45$) than for en-route sectors, as presented in [71, 73] ($R^2 < 0.30$).

5) Workload and Collision Risk: The focus of [76] is on the analysis of predictive capabilities of the complexity indicators (as presented in [4]) towards the collision risk safety indicator. Outstanding to this research is the idea of time-shifting, which introduces a temporal dimension to the correlation between input and output variables. This approach allows to actually treat complexity indicators as precursors for safety-related events. However, using collision risk as safety indicator seems to be a drastic choice, since upstream incidents, like loss of separation, are neglected [77, 78]. In this study, workload is represented by a weighted sum of three components, which are obtainable in a fast-time simulation environment: radio channel utilization time, air traffic control (ATC) idle proportion, and number of deferred tasks. While the publication does not provide the actual model for correlating indicators to workload or collision risk, R^2 values are reported as follows: 0.7 for the workload model, and 0.5 for the collision risk model.

6) Potential Loss of Separation: Another analysis of historic traffic data is presented in [77, 78]. Here, complexity is calculated by NEST. Regarding the quantification of safety, a potential loss of separation (LoS) indicator is considered in addition to the conflict risk. Both metrics are calculated via a customized conflict risk assessment tool [79]. In summary, the analysis found that there is a strong correlation between the daily number of flights and the number of potential LoS ($R^2 = 0.88$), as well as between daily number of flights and conflict risk ($R^2 = 0.80$). For both safety metrics, the correlation with complexity is slightly higher ($R^2 = 0.91$ for PLoS and $R^2 = 0.83$ for CR). There are some more fine-grained results reported in the publication, like the distinction between summer and winter-season, as well as per-flightlevel results. However, besides data analysis, there is no attempt of deriving a correlation model from this. Table I. summarizes the findings.

Table I. Summary of complexity – safety correlation analyses.

<table>
<tr><th>Independent Variables</th><th>Dependent Variable</th><th>R²</th><th>Ref</th></tr>
<tr><td>Traffic Load, Complexity, and Staffing</td><td># of NMACs</td><td>0.61 (TCAs)
0.81 (TRSAs)</td><td>[69]</td></tr>
<tr><td># of Aircraft</td><td>Probability of Operational Error</td><td>0.989</td><td>[70]</td></tr>
<tr><td>Complexity Indic., ISA Ratings</td><td>Propensity</td><td>0.30</td><td>[71, 73]</td></tr>
<tr><td>Complexity Indic.</td><td>Collision Probability</td><td>< 0.45</td><td>[4]</td></tr>
<tr><td>Complexity Indic.</td><td>Collision Risk</td><td>0.5</td><td>[76]</td></tr>
<tr><td rowspan="2">daily number of flights</td><td>Potential LoS</td><td>0.88</td><td rowspan="4">[78, 80]</td></tr>
<tr><td>Conflict Risk</td><td>0.80</td></tr>
<tr><td rowspan="2">NEST complexity metric</td><td>Potential LoS</td><td>0.91</td></tr>
<tr><td>Conflict Risk</td><td>0.83</td></tr>
</table>

V. Conclusion

Already from screening titles and abstracts, but also from full paper reviews, it became clear that the terms *complexity*, *task load*, and *workload* are used inconsistently throughout the literature. Often, these are not clearly distinguished from one another and correlations are assumed implicitly. The majority of publications makes qualitative statements only, which promote the lump-sum assumption of a strictly negative correlation between complexity, or workload and safety.

Quantitative correlation analyses are hardly found in the literature. We consider this lack of formal modeling as a crucial scientific gap: No evidence can be proven to which level workload or complexity can be accepted while granting safety at all operational times. This applies especially to those models that are proposed to give a measure of complexity of the controller task without any further analysis of correlation to any other metric.

Furthermore, it has already been realized that this connection is not as straight-forward as implied. Rather, besides the effects of too high workload, there are also negative effects when workload levels decreasing below a certain threshold ("underload" [81]), which can have compromising impact on safety (e.g. boredom, distraction, unsound work habits, etc.), cf. figure 4 and [82]. For instance, in [64] authors found that lower workload levels are associated with higher frequency of safety events (). Investigating this relationship will become more important with increasing automation beyond a level where support tools generate advice but also implement those, essentially pushing the controller into the passive role of monitoring.

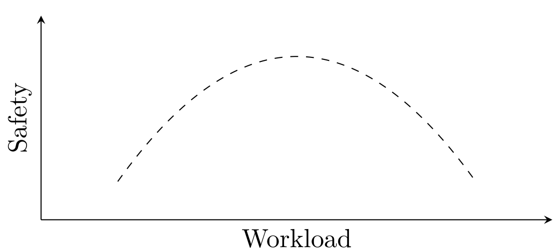

Figure 4. Qualitative depiction of the relationship between workload and safety, according to [4].

From the six models identified that actually investigate and quantify the relationship between traffic- and airspace-complexity indicators and safety metric, we could learn that collision probability might be too specific as such a metric, especially in a predictive context. Corresponding work shows only little correlation between investigated indicators and collision risk [4, 71–73, 75]. Rather, metrics closer to all-day operations of a controller, like loss of separation, are much more suitable, as can be seen in [69, 70, 77, 78].

With recent advances on the field of machine learning and several modern techniques at hand, as well as the increasing availability of large amounts of traffic data (e.g. NASA Sherlock Data Warehouse [83]), we find it somehow surprising that there seems to be no attempt of exploiting these capabilities for investigating the correlation between traffic- and airspace-based indicators and safety metrics. For modelling workload, such approaches have been followed since decades [84–87]. Employing such advanced techniques, we are optimistic that it will be possible to detect patterns in data that no human was able to detect or foresee and which therefore could not be employed for complexity-safety correlation analyses, yet.

The focus of our next steps will be on identifying correlations between potential indicators and safety metrics from historic data with the explicit goal to derive a validated model that can be applied in various contexts, especially for developing the safety-conscious arrival management support tool as highlighted at the beginning.

References

[1] ICAO. "Doc 4444 - Procedures for Air Navigation Services - Air Traffic Management (PANS-ATM)." 16th ed. Amendment 9. Montréal, Québec, Canada, 2020.

[2] R. Mogford et al. "The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature." Technical Note DOT/FAA/CT-TN95/22. Springfield, Virginia, USA, July 1995.

[3] P. N. P. Ferreira et al. "Assessing Operational Impacts of Automation Using Functional Resonance Analysis Method." In: *Cognition, Technology & Work* 21.3 (Aug. 2019). doi: 10.1007/s10111-019-00540-z.

[4] K. Schelbert et al. "Adapting Enroute ATM Complexity Metrics for Terminal Airspace Safety Assessment." In: *5th International Conference on Research in Air Transportation (ICRAT)*. Berkeley, California, USA, May 2012.

[5] A. Liberati et al. "The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Healthcare Interventions: Explanation and Elaboration." In: *British Medical Journal* 339 (July 21, 2009). doi: 10.1136/bmj.b2700.

[6] S. Sporn. "The Air Traffic Controller and Maxwell's Demon." In: *National Aerospace and Electronics Conference (NAECON'94)*. Dayton, Ohio, USA: IEEE, 1994. doi: 10.1109/NAECON.1994.332890.

[7] P. Averty et al. "Evaluating a New Index of Mental Workload in Real ATC Situation Using Psychophysiological Measures." In: *21st Digital Avionics Systems Conference (DASC)*. Vol. 2. Irvine, California, USA: IEEE, 2002. doi: 10.1109/DASC.2002.1052916.

[8] H. Ayaz et al. "Estimation of Cognitive Workload during Simulated Air Traffic Control Using Optical Brain Imaging Sensors." In: *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*. Ed. by D. D. Schmorrow et al. Vol. 6780. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-21852-1_63.

[9] M. Prandini et al. "Toward Air Traffic Complexity Assessment in New Generation Air Traffic Management Systems." In: *IEEE Transactions on Intelligent Transportation Systems* 12.3 (Sept. 2011). doi: 10.1109/TITS.2011.2113175.

[10] B. Juričić et al. "Zagreb Terminal Airspace Capacity Analysis." In: *PROMET - Traffic&Transportation* 23.5 (Jan. 26, 2012). doi: 10.7307/ptt.v23i5.155.

[11] Y. H. Lee et al. "Air Traffic Controllers' Situation Awareness and Workload under Dynamic Air Traffic Situations." In: *Transportation Journal* 51.3 (July 1, 2012). doi: 10.5325/transportationj.51.3.0338.

[12] K. Clegg et al. “The Discovery and Quantification of Risk in High Dimensional Search Spaces.” In: *15th Genetic and Evolutionary Computation Conference (GECCO)*. Amsterdam The Netherlands: ACM, July 6, 2013. DOI: 10.1145/2464576.2464669.

[13] J. Harrison et al. “Human Performance Assessment Study in Aviation Using Functional Near Infrared Spectroscopy.” In: *Foundations of Augmented Cognition*. Ed. by D. D. Schmorrow et al. Red. by D. Hutchison et al. Vol. 8027. Springer Berlin Heidelberg, 2013. DOI: 10.1007/978-3-642-39454-6_46.

[14] W. Z. Tang. “Research of Radar Controllers’ Workload in Fuzzy Synthetic Evaluation Model.” In: *Advanced Materials Research* 722 (July 2013). DOI: 10.4028/www.scientific.net/AMR.722.233.

[15] M. Zhang et al. “Integrated Evluation of the Safety Controller’s Workload Based on Improved Extension Evaluation Model.” In: *Journal of Applied Sciences* 13.15 (July 15, 2013). DOI: 10.3923/jas.2013.2969.2973.

[16] J. Harrison et al. “Cognitive Workload and Learning Assessment During the Implementation of a Next-Generation Air Traffic Control Technology Using Functional Near-Infrared Spectroscopy.” In: *IEEE Transactions on Human-Machine Systems* 44.4 (Aug. 2014). DOI: 10.1109/THMS.2014.2319822.

[17] A. Neal et al. “Development and Validation of a Multilevel Model for Predicting Workload Under Routine and Nonroutine Conditions in an Air Traffic Management Center.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 56.2 (Mar. 2014). DOI: 10.1177/0018720813491283.

[18] B. Számel et al. “Towards Safer Air Traffic: Optimizing ATC Controller Workload by Simulation with Reduced Set of Parameters.” In: *Safety and Reliability: Methodology and Applications*. Ed. by T. Nowakowski et al. London, United Kingdom: Taylor & Francis Group, 2014. DOI: 10.1201/b17399-143.

[19] B. Marr et al. “Controller Workload-based Calculation of Monitor Alert Parameters for En Route Sectors.” In: *15th AIAA Aviation Technology, Integration, and Operations Conference*. Dallas, Texas, USA: American Institute of Aeronautics and Astronautics, June 22, 2015. DOI: 10.2514/6.2015-3178.

[20] C. Surakitbanharn et al. “A Comparison of an Intensity Control Measure Versus Dynamic Density to Capture Complexity Within a Sector.” In: *15th AIAA Aviation Technology, Integration, and Operations Conference*. Dallas, Texas, USA: American Institute of Aeronautics and Astronautics, June 22, 2015. DOI: 10.2514/6.2015-2904.

[21] M. Marchitto et al. “Air Traffic Control: Ocular Metrics Reflect Cognitive Complexity.” In: *International Journal of Industrial Ergonomics* 54 (July 2016). DOI: 10.1016/j.ergon.2016.05.010.

[22] T. Edwards et al. “The Relationship between Workload and Performance in Air Traffic Control.” In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61.1 (Sept. 2017). DOI: 10.1177/1541931213601887.

[23] T. Edwards et al. “Task Demand Variation in Air Traffic Control: Implications for Workload, Fatigue, and Performance.” In: *Advances in Human Aspects of Transportation*. Ed. by N. A. Stanton et al. Vol. 484. Advances in Intelligent Systems and Computing. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-41682-3_8.

[24] T. Edwards et al. “The Relationship Between Workload and Performance in Air Traffic Control: Exploring the Influence of Levels of Automation and Variation in Task Demand.” In: *Human Mental Workload: Models and Applications*. Ed. by L. Longo et al. Vol. 726. Cham: Springer International Publishing, 2017. DOI: 10.1007/978-3-319-61061-0_8.

[25] J. Zhang et al. “A Sector Capacity Assessment Method Based on Airspace Utilization Efficiency.” In: *Journal of Physics: Conference Series* 976 (Feb. 2018). DOI: 10.1088/1742-6596/976/1/012014.

[26] P. A. Dmochowski et al. “A Method of Evaluating Air Traffic Controller Time Workload.” In: *Development of Transport by Telematics*. Ed. by J. Mikulski. Vol. 1049. Cham: Springer International Publishing, 2019. DOI: 10.1007/978-3-030-27547-1_26.

[27] T. Radüntz et al. “The Effect of an Exceptional Event on the Subjectively Experienced Workload of Air Traffic Controllers.” In: *Human Mental Workload: Models and Applications*. Ed. by L. Longo et al. Vol. 1012. Cham: Springer International Publishing, 2019. DOI: 10.1007/978-3-030-14273-5_14.

[28] T. Radüntz et al. “Indexing Mental Workload During Simulated Air Traffic Control Tasks by Means of Dual Frequency Head Maps.” In: *Frontiers in Physiology* 11 (Apr. 21, 2020). DOI: 10.3389/fphys.2020.00300.

[29] V. Socha et al. “Workload Assessment of Air Traffic Controllers.” In: *Transportation Research Procedia* 51 (2020). DOI: 10.1016/j.trpro.2020.11.027.

[30] J. Zhang et al. “The Difficulty to Break a Relational Complexity Network Can Predict Air Traffic Controllers’ Mental Workload and Performance in Conflict Resolution.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 63.2 (Mar. 2021). DOI: 10.1177/0018720819880646.

[31] A. R. John et al. “Unraveling the Physiological Correlates of Mental Workload Variations in Tracking and Collision Prediction Tasks.” In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30 (2022). DOI: 10.1109/TNSRE.2022.3157446.

[32] M. Zamarreño Suárez et al. “How Much Workload Is Workload? A Human Neurophysiological and Affective-Cognitive Performance Measurement Methodology for ATCOs.” In: *Aircraft Engineering and Aerospace Technology* 94.9 (Sept. 9, 2022). DOI: 10.1108/AEAT-11-2021-0328.

[33] Y. Pang et al. “Air Traffic Controller Workload Level Prediction Using Conformalized Dynamical Graph Learning.” In: *Advanced Engineering Informatics* 57 (Aug. 2023). DOI: 10.1016/j.aei.2023.102113.

[34] J. Rebollo et al. “Using Sector Complexity Metrics to Predict Sector Capacity.” In: *42nd Digital Avionics Systems Conference (DASC)*. Barcelona, Spain, Oct. 2023.

[35] H. Wang et al. “Evaluation and Optimization of Air Traffic Complexity Based on Resilience Metrics.” In: *Journal of Advanced Transportation* 2023 (Apr. 4, 2023). Ed. by Y. Du. DOI: 10.1155/2023/5692934.

[36] J. Yang et al. “Cognitive Load Assessment of Air Traffic Controller Based on SCNN-TransE Network Using Speech Data.” In: *Aerospace* 10.7 (June 23, 2023). DOI: 10.3390/aerospace10070584.

[37] B. Li et al. “MAST-GNN: A Multimodal Adaptive Spatio-Temporal Graph Neural Network for Airspace Complexity Prediction.” In: *Transportation Research Part C: Emerging Technologies* 160 (Mar. 2024). DOI: 10.1016/j.trc.2024.104521.

[38] K. Lee et al. “Air Traffic Complexity: An Input-Output Approach.” In: *2007 American Control Conference*. July 2007. DOI: 10.1109/ACC.2007.4282989.

[39] H. Idris et al. “Distributed Trajectory Flexibility Preservation for Traffic Complexity Mitigation.” In: *8th USA/Europe Air Traffic Management R&D Seminar (ATM Seminar)*. Napa, California, USA, June 2009.

[40] M. Prandini et al. “A Probabilistic Approach to Air Traffic Complexity Evaluation.” In: *48th IEEE Conference on Decision and Control (CDC) Held Jointly with 28th Chinese Control Conference*. Dec. 2009. DOI: 10.1109/CDC.2009.5400469.

[41] M. Prandini et al. “Air Traffic Complexity and the Interacting Particle System Method: An Integrated Approach for Collision Risk Estimation.” In: *American Control Conference*. San Francisco, California, USA: IEEE, June 2011. DOI: 10.1109/ACC.2011.5991305.

[42] L. Piroddi et al. “A Geometric Approach to Air Traffic Complexity Evaluation for Strategic Trajectory Management.” In: *49th IEEE Conference on Decision and Control (CDC)*. Atlanta, GA, USA: IEEE, Dec. 2010. DOI: 10.1109/CDC.2010.5718055.

[43] N. Suárez et al. “Quantifying Air Traffic Controller Mental Workload.” In: *4th SESAR Innovation Days (SID)*. Nov. 2014.

[44] H. Wang et al. “Modeling Air Traffic Situation Complexity with a Dynamic Weighted Network Approach.” In: *Journal of Advanced Transportation* 2018 (2018). DOI: 10.1155/2018/5254289.

[45] H. J. Wee et al. “A Spatial, Temporal Complexity Metric for Tactical Air Traffic Control.” In: *Journal of Navigation* 71.5 (Sept. 2018). DOI: 10.1017/S0373463318000255.

[46] M. Zhang et al. “Air Traffic Complexity Evaluation with Novel Complexity Features and mRMR-XGBoost.” In: *IOP Conference Series: Earth and Environmental Science* 638.1 (Feb. 1, 2021). DOI: 10.1088/1755-1315/638/1/012036.

[47] L. Yu Filimonyuk. “Method for Measuring and Analyzing the Complexity and Safety of Air Traffic Control.” In: *Journal of Physics: Conference Series* 2373.5 (Dec. 1, 2022). DOI: 10.1088/1742-6596/2373/5/052019.

[48] W. Du et al. “A Spatiotemporal Hybrid Model for Airspace Complexity Prediction.” In: *IEEE Intelligent Transportation Systems Magazine* (Sept. 2022). DOI: 10.1109/MITS.2022.3204099.

[49] L. Zhang et al. “Air Traffic Complexity Evaluation with Hierarchical Graph Representation Learning.” In: *Aerospace* 10.4 (Apr. 3, 2023). DOI: 10.3390/aerospace10040352.

[50] M. Zhang et al. “Terminal Airspace Sector Capacity Estimation Method Based on the ATC Dynamical Model.” In: *Kybernetes* 45.6 (Jan. 1, 2016). DOI: 10.1108/K-12-2014-0308.

[51] M. Zamarreño Suárez et al. “Methodology for Determining the Event-Based Taskload of an Air Traffic Controller Using Real-Time Simulations.” In: *Aerospace* 10.2 (Feb. 2023). DOI: 10.3390/aerospace10020097.

[52] M. Zamarreño Suárez et al. “Is It Possible to Evaluate the Event-Based Taskload of an Air Traffic Controller Using an Air Traffic Control Simulator?” In: *Journal of Physics: Conference Series* 2526.1 (June 1, 2023). DOI: 10.1088/1742-6596/2526/1/012106.

[53] U. Ahlstrom et al. “Using Eye Movement Activity as a Correlate of Cognitive Workload.” In: *International Journal of Industrial Ergonomics* 36.7 (July 2006). DOI: 10.1016/j.ergon.2006.04.002.

[54] H. Sun et al. “Study on the Development Process of Apron Controller’s Work Fatigue Based on Heart Rate Characteristics.” In: *Heliyon* 10.4 (Feb. 2024). DOI: 10.1016/j.heliyon.2024.e26296.

[55] J. Zhang et al. “A Genetic Algorithm-based BP Neural Network Method for Operational Performance Assessment of ATC Sector.” In: *PROMET - Traffic&Transportation* 28.6 (Dec. 12, 2016). DOI: 10.7307/ptt.v28i6.2003.

[56] Y. Heng et al. “Control–Flight Conflict Interdependent Network Based Controllers’ Workload Prediction Evaluation Study.” In: *Applied Sciences* 12.20 (Oct. 21, 2022). DOI: 10.3390/app122010671.

[57] A. Cloerec et al. “Traffic Complexity Analysis to Evaluate the Potential for Limited Delegation of Separation Assurance to the Cockpit.” In: *18th Digital Avionics Systems Conference (DASC)*. Vol. 1. Oct. 1999. DOI: 10.1109/DASC.1999.863739.

[58] K. Roy et al. “Enroute Airspace Control and Controller Workload Analysis Using a Novel Slot-Based Sector Model.” In: *2006 American Control Conference*. Minneapolis, MN, USA: IEEE, 2006. DOI: 10.1109/ACC.2006.1655497.

[59] F.-q. Dai et al. “Study on the Method of Sector Division Based on the Power Diagram.” In: *International Conference on Measuring Technology and Mechatronics Automation*. Zhangjiajie, Hunan, China: IEEE, 2009. DOI: 10.1109/ICMTMA.2009.151.

[60] T. Koca et al. “A Methodology to Perform Air Traffic Complexity Analysis Based on Spatio-Temporal Regions Constructed Around Aircraft Conflicts.” In: *IEEE Access* 7 (2019). DOI: 10.1109/ACCESS.2019.2928355.

[61] L. Wang et al. “Sector Planning Based on New Complexity Model.” In: *5th International Conference on Transportation Information and Safety (ICTIS)*. Liverpool, United Kingdom: IEEE, July 2019. DOI: 10.1109/ICTIS.2019.8883798.

[62] R. Medianto et al. “Terminal Control Area Complexity Measurement Using Simulation Model.” In: *IIUM Engineering Journal* 24.1 (Jan. 4, 2023). DOI: 10.31436/iiumej.v24i1.2223.

[63] K. S. Lindsay. “Workload-Based Capacity for Air Traffic Management.” In: *Transportation Research Record: Journal of the Transportation Research Board* 2501.1 (Jan. 2015). DOI: 10.3141/2501-02.

[64] T. Edwards et al. “Multifactor Combinations and Associations with Performance in an Air Traffic Control Simulation.” In: *Advances in Human Aspects of Road and Rail Transportation*. Ed. by N. A. Stanton. 1st ed. Boca Raton, Florida, USA: CRC Press, July 9, 2012.

[65] A. A. Mutlu et al. “An Assesment on The Safety and Complexity of The Innovative Design of Istanbul’s New TMA.” In: *9th SESAR Innovation Days (SID)*. Dec. 2019.

[66] S. V. Ligda et al. “Monitoring Human Performance in Real-Time for NAS Safety Prognostics.” In: *AIAA Aviation 2019 Forum*. Dallas, Texas: American Institute of Aeronautics and Astronautics, June 17, 2019. DOI: 10.2514/6.2019-3411.

[67] D. J. Schroeder. “The Loss of Prescribed Separation between Aircraft: How Does It Occur?” In: *Aerospace Congress and Exposition*. Feb. 1, 1982. DOI: 10.4271/821432.

[68] A. Majumdar et al. “Trend Analysis of Controller-Caused Airspace Incidents in New Zealand, 1994-2002.” In: *Transportation Research Record: Journal of the Transportation Research Board* 1888.1 (Jan. 2004). DOI: 10.3141/1888-04.

[69] J. L. Gifford et al. “Airport Congestion and Near-Midair Collisions.” In: *Transportation Research Part A: General* 25.2-3 (Mar. 1991). DOI: 10.1016/0191-2607(91)90128-D.

[70] D. Murphy et al. “A Quantitative Model for En Route Error Rate Analysis.” In: *7th USA/Europe Air Traffic Management R&D Seminar (ATM Seminar)*. Barcelona, Spain, July 2007.

[71] J. Djokic et al. “ATC Complexity as Workload and Safety Driver.” In: *3rd International Conference on Research in Air Transportation (ICRAT)*. Fairfax, VA, USA, June 2008.

[72] J. Djokic et al. “Air Traffic Complexity as a Safety Performance Indicator.” In: *Science & Military Journal* 4 (Jan. 2009).

[73] J. Djokic et al. “Air Traffic Control Complexity as Workload Driver.” In: *Transportation Research Part C - Emerging Technologies* 18 (Dec. 1, 2010). DOI: 10.1016/j.trc.2010.03.005.

[74] EUROCONTROL. “INTEGRA Metrics & Methodologies Safety Metrics Technical Definitions.” Nov. 25, 2000.

[75] M. Vogel et al. “A Quantitative Safety Assessment Tool Based on Aircraft Actual Navigation Performance.” In: *4th International Conference on Research in Air Transportation (ICRAT)*. Budapest, Hungary, June 2010.

[76] M. Vogel et al. “Analysis of Airspace Complexity Factors’ Capability to Predict Workload and Safety Levels in the TMA.” In: *10th USA/Europe Air Traffic Management R&D Seminar (ATM Seminar)*. Chicago, IL, USA, June 10, 2013.

[77] T. Pejovic et al. “Analysis of Relationship between Air Traffic Demand, Safety and Complexity in FABEC Airspace.” In: *9th SESAR Innovation Days (SID)*. Dec. 2019.

[78] T. Pejovic et al. “Relationship between Air Traffic Demand, Safety and Complexity in High-Density Airspace in Europe.” In: *Risk Assessment in Air Traffic Management*. Ed. by Javier Alberto Pérez Castán et al. 2020.

[79] F. Netjasov et al. “Potential Safety Occurrences as Indicators of Air Traffic Management Safety Performance: A Network Based Simulation Model.” In: *Transportation Research Part C: Emerging Technologies* 102 (May 2019). DOI: 10.1016/j.trc.2019.03.026.

[80] CI Consortium. “PJ19.04: Performance Framework (2019).” Deliverable D4.7, Edition: 01.00.01. Nov. 30, 2019.

[81] P. G. A. M. Jorna. “The Human Component of System Validation.” In: *NATO Advanced Study Institute on Verification and Validation of Complex and Integrated Human-Machine Systems*. Ed. by J. A. Wise et al. Vimeiro, Portugal, July 1992. DOI: 10.1007/978-3-662-02933-6_17.

[82] J. W. Danaher. “Human Error in ATC System Operations.” In: *Human Factors* 22.5 (Oct. 1, 1980). DOI: 10.1177/001872088002200503.

[83] M. M. Eshow et al. “Architecture and Capabilities of a Data Warehouse for ATM Research.” In: *33rd Digital Avionics Systems Conference (DASC)*. Colorado Springs, CO, USA: IEEE, Oct. 2014. DOI: 10.1109/DASC.2014.6979418.

[84] G. B. Chatterji et al. “Neural Network Based Air Traffic Controller Workload Prediction.” In: *American Control Conference*. Vol. 4. San Diego, CA, USA: IEEE, June 1999. DOI: 10.1109/ACC.1999.786543.

[85] D. Gianazza et al. “Evaluation of Air Traffic Complexity Metrics Using Neural Networks and Sector Status.” In: *2nd International Conference on Research in Air Transportation (ICRAT)*. Belgrade, Serbia and Montenegro, June 2006.

[86] P. Andraši et al. “Subjective Air Traffic Complexity Estimation Using Artificial Neural Networks.” In: *Promet - Traffic & Transportation* 31.4 (Aug. 10, 2019). DOI: 10.7307/ptt.v31i4.3018.

[87] F. Pérez Moreno et al. “Dynamic Model to Characterise Sectors Using Machine Learning Techniques.” In: *Aircraft Engineering and Aerospace Technology* 94.9 (2022). DOI: 10.1108/AEAT-11-2021-0330.