Relating Air Traffic Controller Perceived Complexity to Characteristics of Individual En-Route Flights

Preliminary Research Report

Amber Stienstra

Technische Universiteit Delft

Relating Air Traffic Controller Perceived Complexity to Characteristics of Individual En-Route Flights

Thesis Report

by

Amber Stienstra

Student number:	4374886	
Thesis committee:	Prof. dr. ir. M. Mulder,	TU Delft
	Dr. ir. M. M. van Paassen,	TU Delft
	Dr. ir. C. Borst,	TU Delft
	Ir. G. de Rooij,	TU Delft
	A. B. Tisza,	EUROCONTROL Maastricht Upper Area Control Centre

Contents

List of Abbreviations

ACC Area Control Centre

AD Aircraft Density

ANSP Air Traffic Service Provider

APW Area Proximity Warning

ATC Air Traffic Control

ATCo Air Traffic Controller

ATM Air Traffic Management

CC Coordinator Controller

CPA Clostest Point of Approach

CPDLC Controller Pilot Data Link Communication

DD Dynamic Density

EC Executive Controller

IO Input-Output

MTCD Medium Term Conflict Detection

MUAC Maastricht Upper Area Control Centre

SESAR Single European Sky ATM Research

STCA Short Term Collision Alert

SUA Special Use Airspace

TBO Trajectory Based Operation

TCT Tactical Controller Tool

VERA Verification and Resolution Advisory

List of Figures

I

Scientific paper

Relating Air Traffic Controller Perceived Complexity to Characteristics of Individual En-Route Flights

Amber Stienstra

***Abstract*—Increasing levels of automation in the ATM system are required to allow for higher traffic densities without exceeding the ATCo's mental capacity. Maastricht Upper Area Control Centre envisions to first automate the control of, so called, "basic" traffic. For this strategy to work, an allocation model is required to determine the complexity of each individual flight entering the airspace, based on objective and predictable flight characteristics, which can allocate a flight to human or automation control. Five metrics are proposed related for the individual flight complexity, and their predictive ability is tested with a human-in-the-loop experiment, where air traffic controllers were tasked with assessing the complexity of an individual flight in a series of static scenarios and indicate which other flights in the scenario affected the complexity score. Results show that the complexity is depends mostly on flights that near each other within 10 NM in the horizontal plane and have overlapping flight levels. Furthermore, flights requiring a flight level change are considered more complex as the number of possible interactions increases due to the vertical change and the additional dimension increases the uncertainty of the trajectory. The performance of the metrics indicated a dependency on different traffic patterns, leading to the conclusion that the use of only one metric is too limited to describe the complexity, and future research is needed to better understand the interdependence between the metrics.**

***Index Terms*—**

I. Introduction

The current Air Traffic Management (ATM) system shows signs of saturation, which lead to delays, rerouting of flights and generally inefficient operations [1]. Further growth in air traffic is limited, in part, by the mental capacity of the air traffic controllers (ATCos) [1, 2]. Steps are currently being taken in the European and American airspace to increase the level of automation, which will aid the ATCos in performing manual and repetitive tasks, reducing the experienced mental load [3]. In Europe these plans are made by the Single European Sky ATM Research (SESAR) Joint Undertaking, a partnership between public and private parties in the ATC sector. More automation and increased connectivity between air traffic service providers should result in improved operation efficiency, allowing higher traffic numbers, without exceeding the ATCos mental capacity.

Increasing automation in the ATC system means that (part of) the tasks, that are currently being performed by the ATCos, will be performed by an automated system. Until a fully reliable system is developed, the ATCo will be responsible for monitoring the performance of this system. Studies into the effects of automation show that humans lack the ability to adequately perform such a monitoring task for a sustained period of time, and that there is a risk of the human losing situation awareness (SA). Also, humans tend to over-trust the automated system, and show a delayed response time in case of an automation failure [4**?**]. To mitigate these negative effects of automation on the human controller, the Maastricht Upper Area Control Centre (MUAC), an Air Navigation Service Provider (ANSP) that provides ATC in the upper flight levels over the Netherlands, Belgium, Luxembourg and part of Germany, proposes a strategy to initially only allocate "basic" traffic to an automated system, while the ATCo is kept engaged with the task of handling the more complex,"non-basic", traffic [5].

A clear definition of basic and complex traffic is needed, however, for this allocation strategy to be feasible [6]. This definition should describe the complexity of an individual flight in terms of observable traffic characteristics. Currently used models for air traffic complexity have mostly been developed in search of a method to find the complexity for the complete system of flights in a defined part of the airspace [7, 8, 9]. These models predict controller workload, and are based on the relation between the objective task load and the experienced controller mental load. The task load is defined by factors such as the total number of flights, the number of crossing routes and the number of climbing or descending flights. The allocation problem posed by the MUAC automation strategy requires a model to determine the complexity not of a system of flight, but instead of an individual flight. Lee et al. proposed a model where the complexity of an individual flight entering the sector is based on heading input needed for the flights already in the sector to keep the system conflict free, only taking into account the current headings of the flights [8]. In this paper the complexity is described from the point of view of the aircraft entering the sector, based on the interactions it will encounter along the planned trajectory.

This paper reviews various metrics for air traffic complexity and proposes a set of flight characteristics that can be be applied in individual flight metrics, which in time may be used to allocate basic traffic to an automated system. A human-in-the-loop experiment was performed to test the predictive validity of these metrics for complexity. This research was done in collaboration with MUAC, who provided the traffic logs for the experiment. The scope of the research is limited to the study of en-route traffic. Real traffic samples were used as a basis for the scenarios to increase the ecological validity of the experiment. These samples were only slightly modified by adding a single flight in different positions, with this controlled manipulation, scenarios with different properties were created.

The structure of the paper is as follows. Section II gives some background information on the expectations for a ATM system with increased levels of automation and the currently used models for traffic complexity. Section III provides a description of factors that are expected to contribute to the complexity of an individual flight. The human-in-the-loop experiment set-up and results in are discussed in Sections IV and V. A discussion of the results and recommendations for future research can be found in Section VI.

II. Background

A. Automation of Basic Traffic

MUAC is developing flight handling system, named ARGOS, that should be able to autonomously perform all ATCo tasks for part of the air traffic [10, 5]. MUAC aims to keep the controller engaged with the task of handling the non-basic, or complex, part of the traffic, while ARGOS handles basic traffic. Communication between ARGOS and the flight deck will be done via Controller Pilot Data Link Communications (CPDLC), a technology that is already available in a part of the commercial aircraft fleet.

For this strategy to work, ARGOS additionally needs to be able to assess the complexity of each individual flight in a traffic situation and allocate the flights to either human or machine control. The allocation will take place when a new aircraft enters the MUAC airspace, thus the complexity prediction should be done before a flight crosses the sector boundary. A sketch of the MUAC proposal can be found in Figure 1, the green flights in the sector are controlled by the human control, the blue flights by the ARGOS system. The black flight at the sector boundary has to be assigned to either human or automation control based on the interactions with other flights along the planned trajectory. The level of responsibility of ARGOS will be increased in three stages. In each of the stages ARGOS will only aid with or control the basic part of the traffic. In the first two stages the ATCo can still take back manual control over a flight after it has been allocated to ARGOS. In the third, and final, stage ARGOS will autonomously control basic traffic, performing all ATC tasks. In this stage the ATCo will no longer be responsible for the ARGOS traffic and will not be monitoring these flights. An option will be available, however, for ARGOS to indicate to the ATCo that supervision is required.

In the traffic picture resulting from this strategy, flights controlled by a human controller and by automation will travel in the same airspace. In an exploratory study into this mixed responsibility concept ATCos indicated that, apart from the traffic complexity, the possibility of an interaction between traffic controlled by a human controller and the automation should be considered in the allocation scheme [6]. The ATCos argued that a single automated flight interacting with a number of manually handled flights would add unwanted uncertainty to the traffic situation. Other factors that were found to be important for the allocation strategy are the type of flight (overflight, departure, etc.) and the capabilities of the automated system. These factors may be useful in a later stage of the ARGOS project to tune the allocation algorithm. This paper focuses on the search for factors that can describe the complexity of an individual flight, and the use of these factors to determine whether a flight is basic or non-basic.

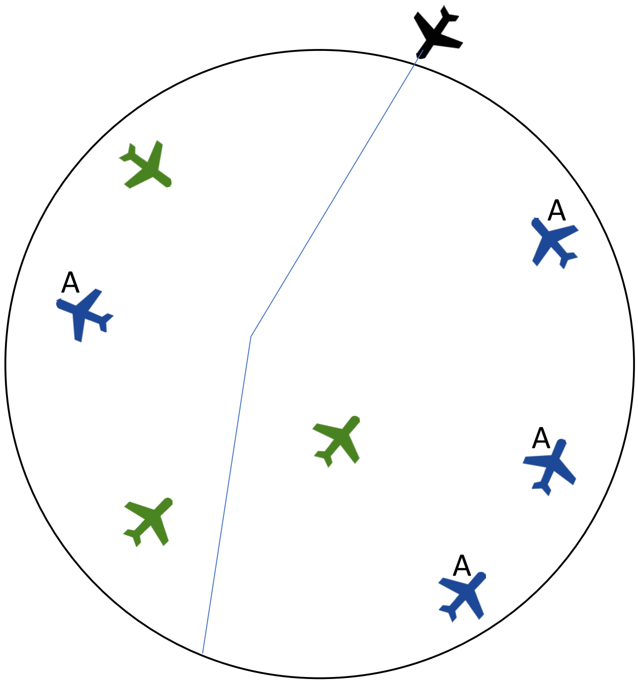


Fig. 1: Scenario sketch of the MUAC automation strategy. Automated flights indicated by the blue colour and A, human controlled flights are green. The flights at the sector entry still has to be assigned.

B. Current Traffic Complexity Metrics

Metrics that describe air traffic complexity were developed in the search for a model for the controller mental load as function of the task load [14]. The task load is the observable difficulty of the controller task, in terms of airspace and traffic characteristics and the concept of operation. The mental load actually experienced by the ATCo is affected by factors such as their skill level, experience, strategy, and mental state [**?** 14]. To quantify the difficulty of the task several air traffic complexity metrics have been developed. Table I gives an overview of commonly used metrics from current air traffic complexity models.

The metric are currently used for workload predictions. For example, to find the complexity of the traffic in a sector, the Dynamic Density (DD) metric combines part the metrics in Table I, describing the complexity as a combination of factors related to the sector geometry and interactions between the traffic [7]. For the DD metric the relative effect on the complexity has to be determined for each of the metrics used, which is done using subjective ratings from ATCos for different traffic scenarios, making the DD metric sector dependent. Sector dependence is undesirable for this project, as it would require the construction of a unique allocation algorithm for each sector. However, a model of the complexity of an individual flight will also take into account different metrics, for which the relative effect on the complexity has to be found, in this process it is important to avoid creating a sector dependent model. Therefore, for this project different sectors are used in the experiment.

The metrics for sector complexity, in Table I, illustrate how traffic characteristics contribute to the controller experienced complexity. Currently, these metrics are mainly used for instantaneous workload prediction, counting the occurrence of a situation at a given time. For the allocation of a flight to ARGOS a prediction of the complexity along the entire

TABLE I: Traffic complexity metrics identified from literature

Metric	Description	Relevance for individual flight complexity
Aircraft count	The total number of flights in the airspace. The number of flights is positively related with the sector complexity.[11]	Not all flights in the sector are relevant for the complexity of a single flight as flight. Instead the number of flights nearing the individual flight can be considered.
Number of aircraft/volume of airspace	This metric indicates the available space for each aircraft in the sector. Less space available per flight leads to higher sector complexity [7].	To assess the complexity of an individual flight the number of of flights crossing it's trajectory can be considered. The moment of the crossing can be used to determine of a flight interacts with a cluster of flights, in a short time frame, or if the crossings are spaced out along the trajectory.
Proximity of conflicting aircraft with respect to their separation standards	Closeness of flights to their separation standards at their closest point of approach. If flights can violate the separation minima in the vertical and lateral plane, a conflict only occurs when separation is lost in both dimensions simultaneously. The experienced complexity increases when two flights approach the violation minima in one or both dimensions [7].	Instead of the closeness to the violation minima of all flights, the closeness of a all flights to the individual flight can be considered.
Number of climbing or descending aircraft	A count of the number of climbing or descending aircraft at an instant in time. Climbing or descending flights add to the complexity as the change in flight level might lead to new conflicts in the traffic scenario, and requires the ATCo to extend his/her mental representation of the traffic scenario [7].	This metric is relevant to the complexity of an individual flight in two ways. First, the required vertical change of the individual flight adds to the complexity. Second, if the flight interacts with other climbing or descending flights this adds another dimension to the trajectory predictions, thus this metric also affects the complexity of the interactions of the individual flight [12].
Time-to-go to loss of separation	The amount of time remaining until the violation of separation minima. A shorter time to the moment of loss of separation add to complexity as there is less time for the ATCo to resolve conflict [7].	This metric can be applied to assess the complexity of the interactions of the individual flight. For each potential conflict situation of a flight the time until the separation limits are reached can be calculated.
Ratio of standard deviation of speed to average speed	This metric is most relevant for traffic patterns with aircraft streams, as speed deviation causes bunching in the airspace. Furthermore, research shows that more cognitive effort is needed to determine whether two flights on converging tracks will get into a conflict if the flights are travelling at different speeds [7, 13].	Similar to the time-to-go to loss of separation the relative speed for a potential conflict can be calculated.
Conflict resolution difficulty based on crossing angle	The detection difficulty of a conflict is affected by the conflict angle between the crossing flights. For a 90 degrees crossing angle detection is easiest, sharper crossing angles increase the conflict solving effort [7, 14, 15].	This metric can be applied to assess the complexity of the interactions of the individual flight by finding the crossing angles of the interaction.
Count of number of aircraft within a threshold distance to the sector boundary	The proximity to the sector boundary limits the resolution options for a conflict, increasing the complexity of the resolution [7].	A potential conflict occurring near the sector boundary may be considered more complex as there is a more limited solution space. Instead of increasing the complexity score for interaction near the sector boundary, a metric can be defined that describes difficulty of finding a resolution in terms of the solution space around the flight, considering the effects of both the sector boundary and the other flights.

trajectory in the sector has to be made, before the flight enters the sector. The prediction has to be based on the interactions that are predicted to occur along the trajectory of each individual flight, therefore, metrics related to the sector geometry, such as the proximity to the sector boundary and sector size, are less relevant for this prediction.

III. Complexity Factors for an Individual Flight

The metrics for air traffic complexity discussed in Section II-B show that the complexity of the system of flights depends also on how flights interact in the airspace. Complexity increases if a flight interacts with other flights along its trajectory, different properties of these interactions further affect the complexity. As the ATCo is tasked with solving potential conflict situations, the ease of finding a resolution affects the experienced complexity, which also depends on the distribution of the flights and other factors that limit the flight's movement, such as the sector boundary or trajectory requirements.

From the metrics for sector complexity five metrics, are derived that describe the complexity of an individual flight. The metrics describe the complexity of the flight from the perspective of a single flight, taking into account the flight's trajectory and exit flight level (XFL). To take into account the effect of complexity of finding a resolution for a conflict on the trajectory, these five metrics can then be used to describe any other trajectory and assess the difficulty of finding a conflict-free trajectory. For the experiment only the current trajectory and the direct trajectory to the sector exit point are considered. The metrics are defined as follows:

- **Number of potential loss of separation (LOS) situations on the trajectory**
  A LOS situation occurs when a flight comes within 7 NM of each other without sufficient vertical separation of the considered. Discussions with MUAC ATCos revealed that action will always be taken if this occurs [1]. The output of this metric is the number of predicted LOS

[1]Personal communication on 14-02-2022

situations on the trajectory.

- **Number of crossings on the trajectory**
A crossing occurs when a flight comes within 10 NM of the considered flight in the lateral plane. A crossing does not always require action from the ATCo, as there is no violation of the separation minima. However, additional attention is given by the ATCo to flights that approach each other in the lateral plane. The amount of attention depends on the effort needed to determine whether loss of separation could occur, for which the vertical trajectory of the two flights has to be considered [16]. A high number of crossings adds to the visual clutter on the radar screen, which can increase complexity [17]. Lastly, this metric is an indicator of the impact on the sector of changing the trajectory of the considered flight. Changing the trajectory of a flight with a high number of crossings, means that the ATCo has to re-consider the effects of the change for each of those crossings. The output of this metric is the predicted number of crossings on the trajectory.
- **Closest point of approach (dCPA) [NM]**
The closest point of approach is the minimum distance in the lateral plane that will occur between the considered flight and any other flight in the sector. An additional distinction for this factor can be made between flights with or without sufficient vertical separation. In a discussion about complexity, before the experiment, MUAC ATCos indicated that the dCPA is the most influential factors regarding complexity. If the minimum distance is close to, but not below, the violation limit, action is usually taken to ensure separation is maintained. MUAC ATCos indicated that they take action in case the predicted minimum distance is below 7 NM. For a distance between 7 and 10 NM the ATCos allocate additional attention to the flight.
- **Time to the loss of separation moment (tLOS) [s]**
The tLOS is the time remaining until loss of separation occurs between the considered flight and any other flight in the sector. This limits the time available for an ATCo to find a solution [18]. For the detection of possible conflicts ATCos make predictions about the future states of the flights in their sector. The time over which this is done, the look-ahead time, is affected by the sector geometry. Therefore, the effect of the temporal proximity to the loss of separation moment on complexity may be different per sector. Despite these differences MUAC ATCos did indicate that any situation that requires action, that is unsolved and less than three minutes away generally makes them nervous. Figure 2 shows the construction of the dCPA and tLOS metric.
- **Required flight level change (FL)[ft]**
The required flight level change is the difference between the current flight level of the flight and the required flight level of the considered flight. A flight level change is usually required for flights close to their departure or arrival airport, but over-flight traffic may also request a flight level change, for operational purposes. Making prediction for the futures states of climbing or descending traffic requires additional cognitive effort from the ATCo and leads to a higher level of uncertainty, making a flight more complex [16].

Figure 3 and Table II show the outcomes of the metrics for a simple scenario with three flights. The metrics are applied to the flight with callsign CT22C, both the current trajectory and direct (DCT) trajectory are shown in the Figure. The metric shows that the dCPA between CT22C and WAW49 changes from 3.8 NM to 7.2 NM if CT22C flies direct to the sector exit point. The dCPA and tLOS information of all pairs is used to determine the number of LOS situations and crossings. The tLOS and dCPA can also be used to describe the geometry of a single interaction, which can be analysed separately for their effect on the perceived complexity.

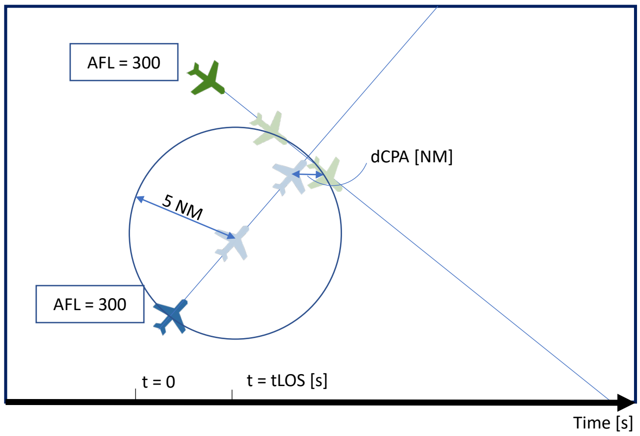


Fig. 2: Construction of the CPA and tLOS metric.

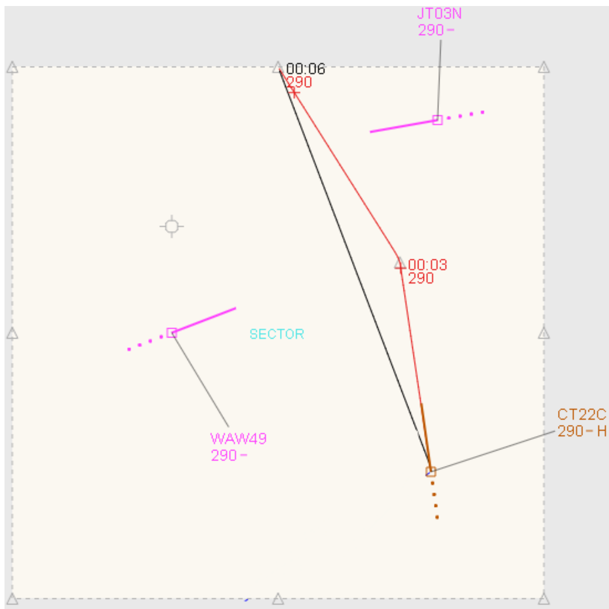


Fig. 3: Screenshot of an example scenario for the metric outcomes with inverted colours. For CT22C the current trajectory and DCT trajectory are visible, the DCT trajectory is a direct path to the sector exit point.

IV. Method

The the experiment aimed to determine the ability of individual flight complexity metrics to predict the experienced complexity.

A. Scenario Design

Three sets of static scenarios were designed, based on three MUAC sectors: Brussels, Jever and Munster. The basis for

TABLE II: Metric output for the CT22C in the example scenario in Figure 3

Metric	Current trajectory	DCT trajectory
FL change [ft] (CT22C)	0	0
number of LOS situations	1	0
Number of crossings	1	1
WAW49		
dCPA [NM]	3.8	7.2
tLOS [s]	130	n/a
JT03N		
dCPA [NM]	14.8	11.6
tLOS [s]	n/a	n/a

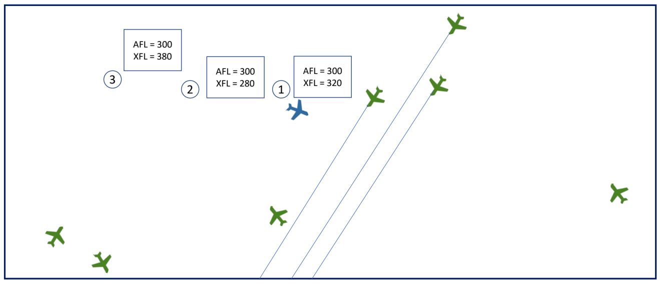


Fig. 4: Scenario design: Flight level change.

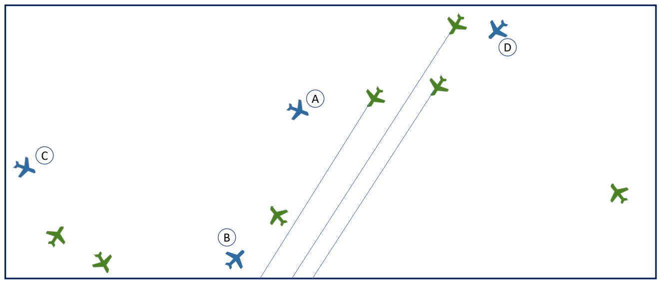


Fig. 5: Scenario design: Multiple sets of scenarios.

each set of scenarios was the same moderately busy traffic sample, obtained from flight plan and radar data from MUAC of a six hour sample of a nominal day. Three sectors were used to investigate the robustness of the proposed metrics and to avoid a sector dependent outcome. As background traffic positions and sector layout were unique for each sector, also the total number of flights in the background traffic varied between the sectors (Brussels = 32 flights, Jever = 23 flights, Munster = 15 flights).

A single flight, the flight of interest (FOI), was added to each scenario. By changing the position, flight level and direction of the FOI, interactions with varying properties were created. Scenarios were created in groups of three using the following methods:

Set 1 - tLOS (time to the LOS moment)

- Independent variable: The time to loss of separation was varied to between (approximately) 300, 500 and 800 s.
- Control variables: The interaction geometry, in terms of crossing angle and CPA were kept constant. The position of the FOI was varied to have a different value of time left to the conflict.

Set 2 - DCT (no interactions DCT path)

- Independent variable: For each scenario the FOI was cleared on a trajectory on which a conflict would occur if no action was taken; this conflict could be solved by either clearing the FOI direct to the sector exit point (COPX) or by clearing the FOI to the XFL.
- Control variables: The screen position of the FOI, COPX and XFL remained unchanged.

Set 3 - FL (flight level change of FOI)

- Independent variable: The required flight level change of the FOI was varied between a small climb or descent (2,000-4,000 ft) or a large climb or descent (>4,000 ft).
- Control variables: The position and direction of the FOI remained unchanged to isolate the effect of the required flight level change on the complexity assessment.

Figure 4 illustrates the design method of the flight level change scenarios. The flight labels of the FOI (1,2,3) indicate the three options for the required level change.

To increase the number of scenarios, for each set multiple groups of scenarios were created. This process is illustrated in Figure 5, four possible screen positions were used (A,B,C,D) at each of these positions three scenarios were created. Figure 6 provides a schematic overview of all scenarios.

The metric introduced in Section III describe the complexity of the FOI for each scenario based the following factors: the number of LOS situations of the FOI, the number of crossings of the FOI. the dCPA of the first interaction [NM], the tLOS of the first interaction [s], and the required flight level change of the FOI [ft].

B. Participants and Apparatus

Fourteen professional MUAC ATCos participated in the experiment. The trials were scheduled during office days for the ATCos, when the ATCos perform tasks other than ATC, or at the end of their ATC shift. The ATCos at MUAC can be licensed for one of three sector groups, that each cover a part of the MUAC airspace. The participating ATCos were only presented with scenarios of their own sector group during the experiment. Table III shows the division of ATCos between the sector groups and other characteristics.

During the experiment the standard MUAC radar screen was mimicked using a Java based simulator, which could be controlled with a computer mouse and keyboard inputs.

TABLE III: Characteristics of the participants

	Brussels	Jever	Munster
Number of participants	4 (all male)	5 (4 male, 1 female)	5 (all male)
Age, years (std)	38.5 (3.5)	40.8 (7.6)	42 (5.4)
Experience, years (std)	15 (0)	15.4 (7.2)	19.4 (6)

C. Procedure

The experiment started with a short briefing and training phase. During the briefing the ATCos were told that the goal of the experiment was to assess the complexity of an individual flight (the FOI). For each scenario their job was to focus on the effort it would require to bring the FOI from the current location to the required sector exit point, and to indicate what flights from the background traffic they considered in their complexity assessment.

With a practice scenario, containing only two flights in an artificial square sector the controls of the simulator were explained. Next, each ATCo was given four training sessions, in which the background traffic and sector were the same as in the experiment scenarios, allowing them to familiarise themselves with the controls and the background traffic.

During the experiment, the ATCos were presented with a static scenario, where the flights did not move. Scenarios from the three sets were given in a partially randomised

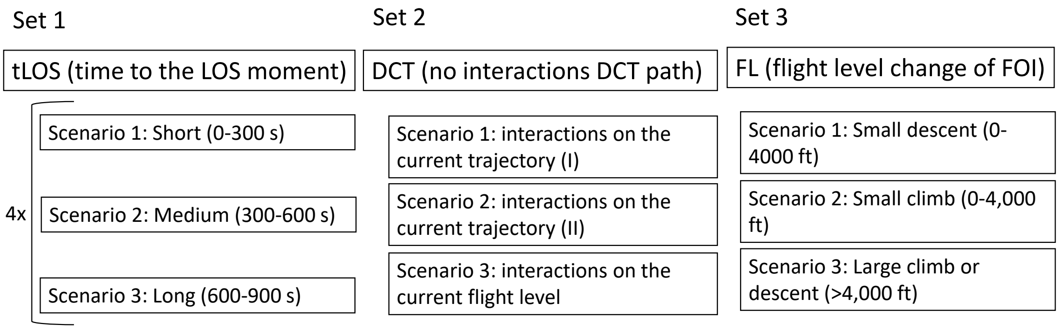
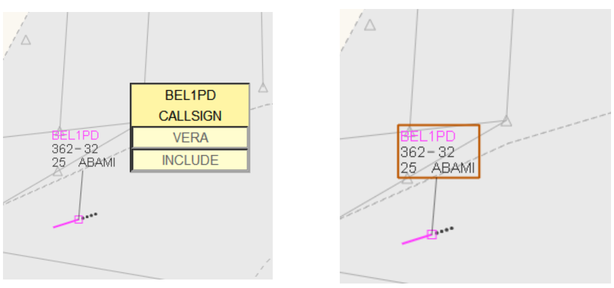


Fig. 6: Overview of all scenarios.

order, where no scenarios from the same subset of three scenarios (Figure 6) would follow each other. By hovering over the flight labels the ATCos could see the sector exit point, heading, altitude and velocity of the flight, by clicking on the flight label they could see the currently planned trajectory. Furthermore, the VERA (Verification and Resolution Advisory) tool was available, a MUAC tool to see a prediction of the horizontal dCPA between two flights, as were the velocity leaders, that show where each flight will be after a fixed amount of time, to assess the complexity of a scenario.

For each scenario, the participants first had to indicate flights that they considered in their complexity assessment, by selecting the 'include' option in the flight label menu. This process is illustrated in Figure 7. After pressing a button to indicate that they finished including flights, they were asked to rate the complexity of the FOI by clicking on a 0-100 scale. Figure 9 shows the radar screen and complexity rating scale used during the experiment. Figure 8 shows the general planning for the experiment; as there was no time limit for the complexity assessment the total time spent on the scenarios varied between the participants.

After rating the scenarios, the ATCos were asked to fill out a questionnaire about nine scenarios. Out of each set of twelve scenario the highest, lowest and middlemost previously rated were reviewed to gain more insight into the reasoning behind the reported complexity. The scenarios were reviewed in the same order as they were presented in the experiment. The participants could see the complexity score they gave, but did not have information on the scenario types or which scenarios were taken for the review.



(a) Flight include menu. (b) Indication of included flight.

Fig. 7: Flight inclusion process.

D. Dependent Measures

To assess the correlation between the single flight complexity metrics and the experienced complexity of a flight the following measures were defined.

- *Subjective complexity rating*
  The participants assessed the complexity of each scenario on a 0-100 scale. To compare the complexity score between participants, the measurements were corrected by calculating the Z-scores for each participant, using the following equation:

$$S_{Z_{i,j}} = \frac{S_{i,j} - \bar{S}_i}{\sigma_i} \tag{1}$$

  In this equation $\bar{S}_i$ and σ_i represent the participant's mean score and standard deviation, respectively.
- *Included flight list*
  For each scenario the participants were asked to indicate which flights from the background traffic they considered in their complexity assessment. This list of flights was used to analyse the correlation between interaction geometries and the experienced complexity.
- *Questionnaire responses*
  The responses of the questionnaire were used to gain further insight into the reasoning for the given complexity ratings. In the questionnaire, the ATCos were asked to indicate factors that contributed to their complexity assessment. The responses were related to the following topics:
  - What factors contributed to the complexity score of the FOI? (e.g., a presence of a loss of separation situation, the fact that a flight was climbing or descending, or the solution space around a flight.)
  - For each included flight: What factors did you take into account when including this flight?
  - How comfortable would you be with the FOI being handled by automation?

E. Hypotheses

For the metric output and the perceived complexity the following relations were hypothesised:

HP 1 The number of LOS situations is positively correlated with the perceived complexity

HP 2 The number of crossing is positively correlated with the perceived complexity

HP 3 The dCPA is negatively correlated with the perceived complexity

HP 4 The tLOS is negatively correlated with the perceived complexity

HP 5 The required flight level change is positively correlated with the perceived complexity

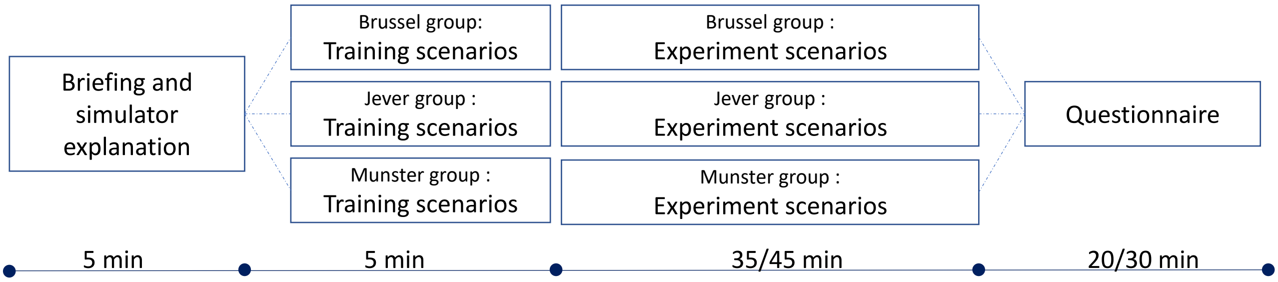


Fig. 8: Planning of the experiment

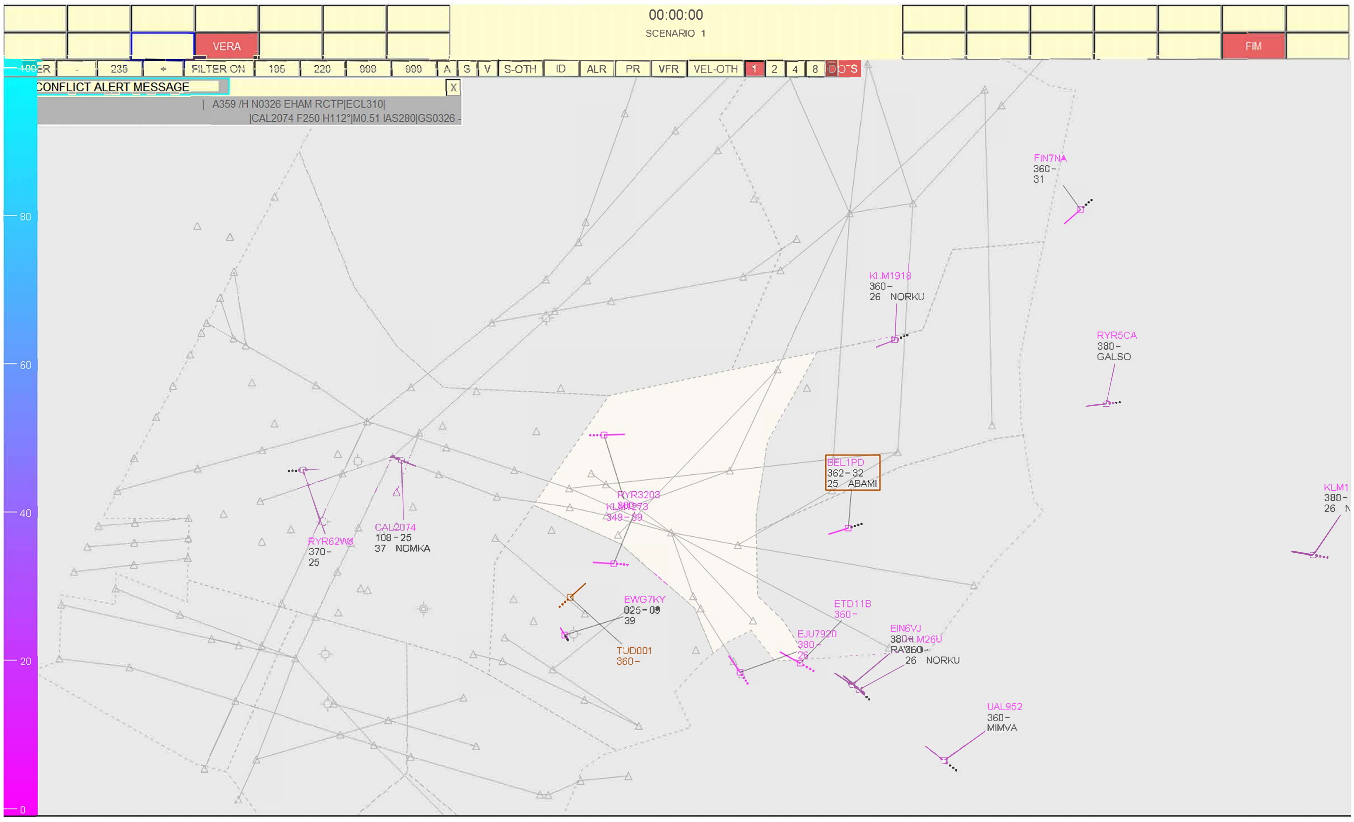


Fig. 9: Screenshot of experiment set-up with inverted colours for clarity. The radar screen with similar lay-out to the screen used at MUAC. The white area is the Munster sector, for which the ATCo was responsible. The surrounding grey area is outside of the ATCo responsibility. The complexity score could be given by clicking the bar on the left. The FOI is shows in orange (TUD001), all pink flights are the background traffic.

V. Results

All fourteen participants finished the experiment. In the questionnaire ten ATCos indicated that they found the traffic sample moderately realistic, three others said it was very realistic, one ATCo found the scenarios slightly realistic. Also the simulator environment was rated to be very realistic by six of the ATCos and was moderately realistic according to the other eight.

The results section starts with an analysis of the complexity scores given for the different sector and scenario types. Next, the relation between the number and properties, in terms of proximity to the FOI, of the included flights are discussed. Some of the outcomes of the included flight analysis are also confirmed in the analysis of the individual metrics. Finally, to gain more insight in the factors that the ATCos considered in their complexity analysis the questionnaire results are presented.

A. *Complexity Score Analysis*

1) *Complexity score variance:*

The Z-score from each participant per scenario was compared to assess the level of consensus on the complexity between the different scenarios. Figure 10 shows the Z-score for all 36 scenarios from each participant of the Brussels sector, the scenarios are ordered by the spread in complexity score. There is a clear variance in the given scores, indicating that the experienced complexity indeed varied between the scenarios. This same variance is visible in the results for the Jever and Munster sectors, these plots can be found in the appendix+. For the majority of the scenarios there is a large

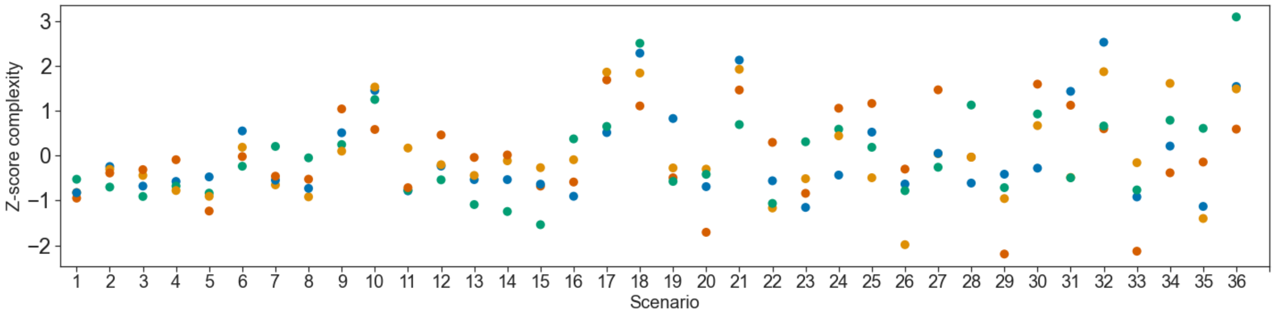


Fig. 10: Z-scores of Brussels participants per scenario ordered by spread in complexity score. Each column of scores represents one of the 36 scenarios. The scores for each participants are given in a different colour.

spread in the scores given by the participants. This differences can come from differences in complexity scoring strategies. In other cases there are differences in working strategies, for instance in which interactions an ATCo thinks require a resolution.

2) Sector complexity:

Apart from the between-participant variance in complexity per scenario, there is also a difference in the ranges of the scores given per sector. Figure 11 shows all complexity scores given per sector. Scores for the Jever and Brussels sectors are given within a similar range, while the range of the Munster scores is a little smaller, with the highest complexity scores being around 20 points lower than the Jever and Brussels scenarios. This difference may be explained by the fact that the number of flights in the Munster scenarios was lower than in the other two sectors. Still, the bulk of the complexity scores for all sectors are concentrated around the same range. The Kruskal-Wallis test showed that the variance in the scores between the different three sectors was significant (H(2)=21.6, p=2.02e-05). Further pairwise comparison showed that the difference between the Brussels and Munster sector (p=0.000012) and the Jever and Munster sector (p=0.018826) was significant.

3) Detection of LOS situations:

Most scenarios were designed such that the FOI would have an interaction on its current trajectory within 900 s from the starting positions of the scenario. Only one of these LOS situations was not included by one participant during the experiment. When presented with this scenario in the review-phase, the ATCo realised, on his own, that he had failed to included it. Despite this miss, the associated complexity score does not seem to be an outlier, it is within the range of scores given by this participant and it is not the lowest score given for this scenario. It is possible that the ATCo did register the interaction, but simply failed to include it.

4) Scenario type analysis:

Three different sets of scenarios were used in the experiment, each with different characteristics. Figure 12 shows the boxplots of Z-scores from all participants of all sectors per scenario type. Despite the variation in complexity scores given between the individual scenarios, on average the FL scenarios received the highest scores. Thus, a required flight level change for the FOI increased the experienced complexity. In the DCT scenario, a pending conflict could be solved by letting the FOI go direct to the XCOP or to the XFL, which means that part of the DCT scenarios had a required FL change. The option for this "easy" resolution did not result in a significantly lower complexity score however, which may be in part because of the effect of the required FL change. In fact, the tLOS scenarios, where the FOI was already on a direct path to the XCOP, were rated lowest out of the three scenario types, despite an interaction being present on the direct track. The Friedman test showed that the difference in Z-score, reported for the three types of scenarios, was significant only for the Jever sector ($\chi^2(2)$ = 8.63, p = 0.01), Pairwise comparison with the Wilcoxon signed-rank test with Bonferroni correction shows that only the difference between the tLOS and FL scenarios is significant (p=0.004488).

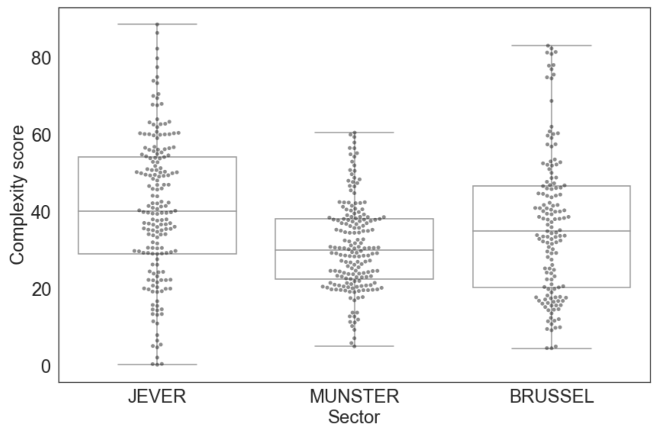


Fig. 11: All complexity scores given per sector from all participants.

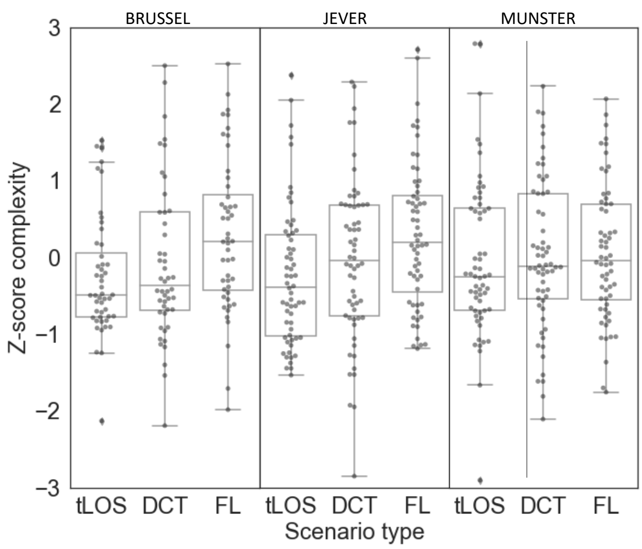


Fig. 12: Z-scores reported per scenario type and sector.

B. Included Flights Analysis

Figure 15 shows a positive relation between the number of included flights and the reported complexity for all sectors. Thus, in case a higher number of flights is included the complexity of the FOI is reported as higher. The number of flights included varied between the participants. For example, for the Brussels scenarios Participant 4 included a maximum of six flights, while Participant 3 included up to 13 flights. Still, both average complexity lines show a similar positive trend. The flights that were included are not solely potential conflict situations; the ATCos indicated many other flights that also affected the complexity assessment.

Using the Kendall-Tau test this positive correlation was found to be moderate but significant for all sectors (Brussels R=0.49 $p<0.001$, Jever R=0.54 $p<0.001$, Munster R=0.53 $p<0.001$).

The remainder of this section discusses the further analysis of included flight's relation to the FOI. For this analysis the proximity, in terms of dCPA and tLOS in the horizontal plane, of each flight to the FOI was described by dividing the flights in six categories, as shown in Figure 13. Only flights with a dCPA under 10 NM and a tLOS under 900 s were considered. From the 828 included flights only 21 flights were outside this range, they were not considered as they lie in a very wide range around the six categories and make up only around 1% of the total flights in that range.

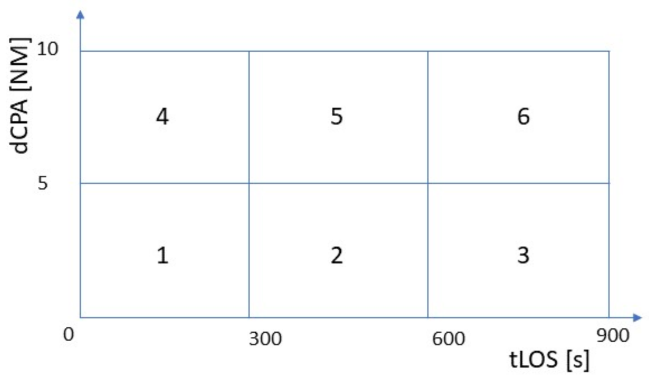


Fig. 13: Proximity to FOI categories in terms of tLOS and dCPA.

The flights were also categorised by their flight level band, which was constructed from the flight's current flight level and the XFL. Flights with flight level bands that intersect with the flight level band of the FOI were marked as intersecting.

In Figure 16 all flights in the scenarios are plotted by their proximity to the FOI, different bars are used for intersecting flights and flights without intersecting flight levels. Each bar also shows what percentage of the flights in each category was included by an ATCo in the experiment. The number of flights in Category 4 is relatively low as no this type of interaction was not created in the design of the scenarios.

The most notable result from this analysis is that hardly any flights were included that did not intersect the FOI's flight levels, for all sectors combined only 28 out of 828 included flights. For each of the sectors most (or all) of the flights in the first proximity category were included. For the Jever and Munster sectors, a large portion of the flights in the second and third category, where the temporal proximity to the CPA moment is larger, were also included. This result is slightly

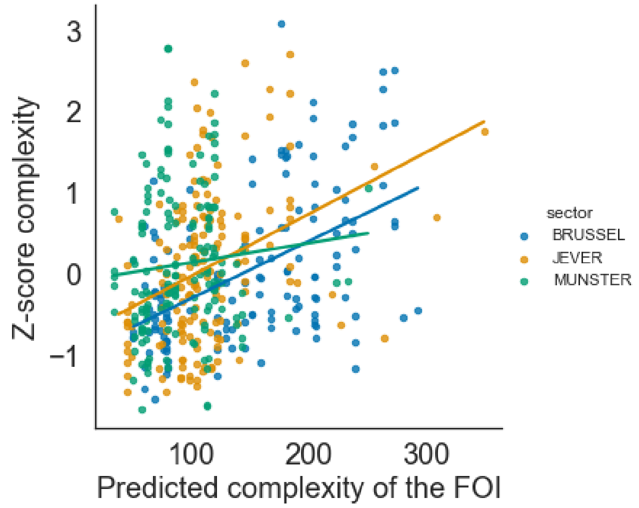


Fig. 14: The predicted complexity of the FOI, based on the percentage of flights included in each separation category, against the reported complexity.

different for the Brussels sector, where just over half of these flights were included. Interestingly, for all sectors also flights have been included from the fourth, fifth and sixth categories. Apparently, despite having a predicted dCPA larger than the horizontal violation limit they still add to the experienced complexity.

The percentage of the included flights in each category is used to make a prediction, in a first attempt to create a model of the complexity of the FOI. Each flight in a scenario is given a complexity score based on the average percentage of flights included in this category during the experiment. For example a flight in Category 1 gets a score of 9.1 (((93% + 79% + 100%)/3)/10). The predicted complexity is plotted against the reported complexity in Figure 14. The Kendall-Tau test shows that the correlation between the predicted complexity and the reported complexity is weak but significant for all sectors (Brussels R=0.33 $p<0.001$, Jever R=0.27 $p<0.001$, Munster R=0.11 $p=0.04$).

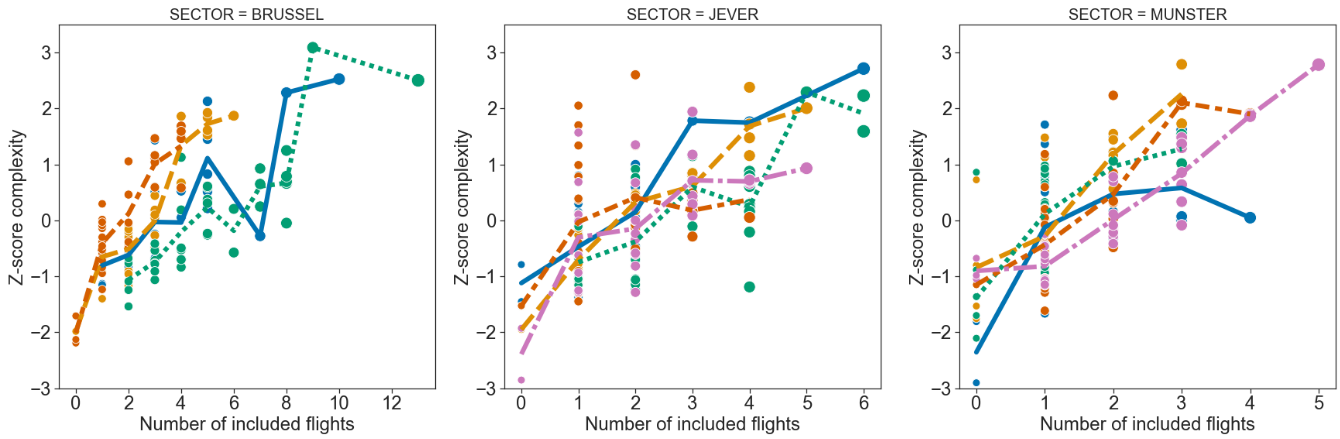


Fig. 15: Number of included flights/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant.

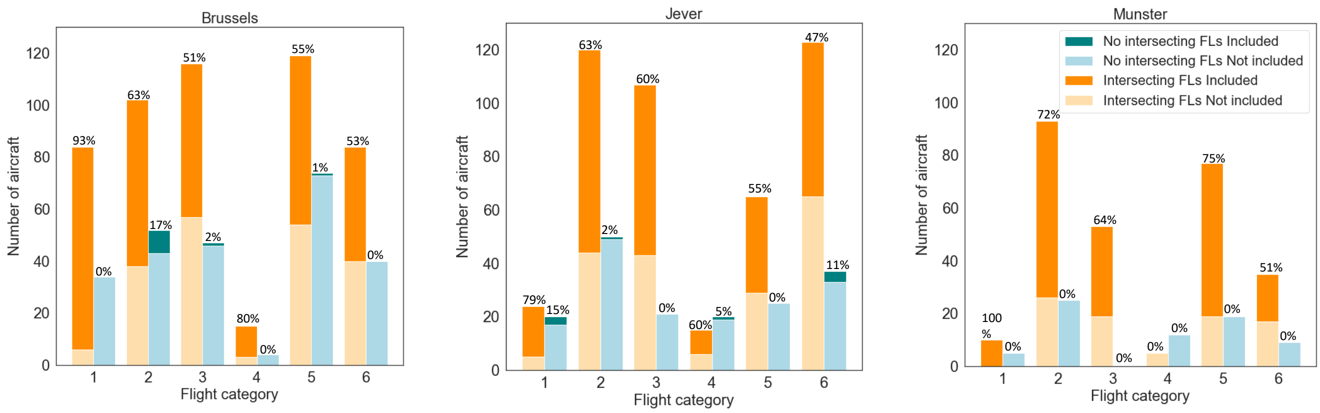


Fig. 16: Total number of flights in each separation category and the percentage of flights included during the experiment

C. Metric Analysis

The outcomes of each metric and the corresponding Z-score per participants for the Brussels and Munster sectors are plotted in Figures 17 to 21. The most significant correlations were found for the Brussels sector, the Jever and Munster sector showed a similar outcomes, therefore only Brussels and Munster are plotted here. Plots for the Jever sector can be found in the appendix. The lines in these figures are the average complexity Z-score given by a participants for scenarios with the same metric outcome. Since most data were not not normally distributed, a Kendall-Tau test was performed to determine if a significant correlation exist between the metrics and the Z-scores. Table IV shows the correlation results between the five metrics and the Z-score for all three sectors, with R indicating the strength of the relation between the two variables.

1) number of LOS situations: The average score lines in Figure 17 show a positive trend between the number of LOS situations and the given Z-score of the Brussels sector. A weak but significant positive corresponding correlation was found, which is the strongest out of all the metrics. For the Munster sector the number of scenarios with two LOS situations was relatively low, making it difficult to identify a similar relation.

2) Number of crossings: Figure 18 shows an upward trend between the number of crossings in the scenario and the Z-score only for the Munster sector, this trend is a weak but

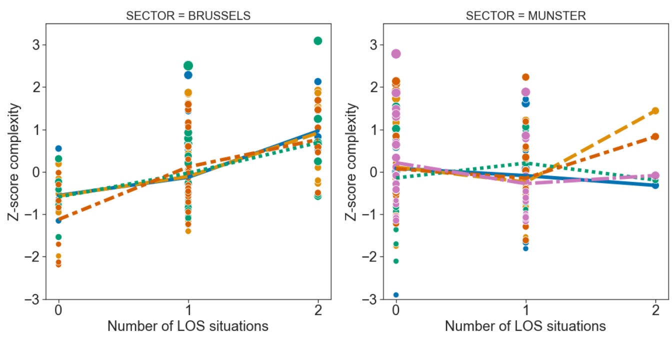


Fig. 17: number of LOS situations/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant. The size of the dots indicates the number of included flights in the scenario.

significant correlation. For the Brussels sector a positive trend is visible until 12 crossings, after which the trend becomes negative. This metric shows a large spread in scores given, even by one participant, for different scenarios with the same number of crossings, indicating that other factors influenced the complexity score.

3) dCPA: Figure 19 shows a negative trend between this metric and the Z-score for the Brussels scenarios. The scores of one Brussels participant for the scenarios with a dCPA of 7-10 NM are significantly lower than for the other

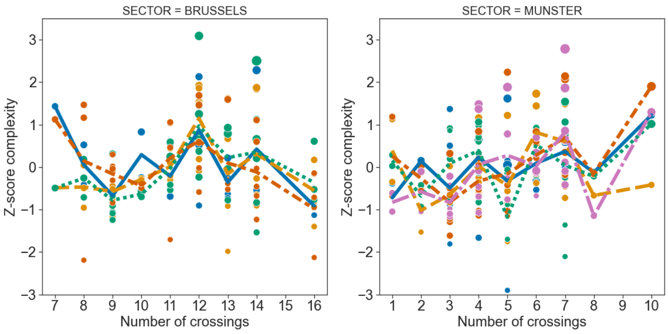


Fig. 18: Number of crossings/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant. The size of the dots indicates the number of included flights in the scenario.

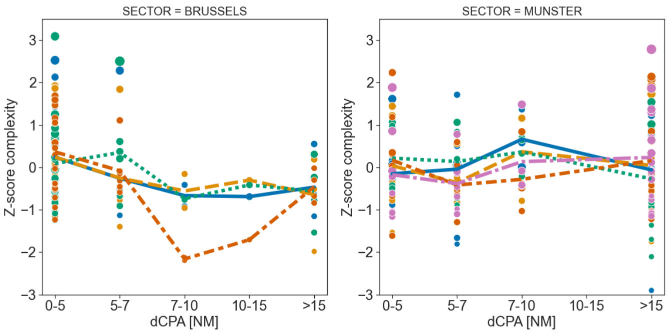


Fig. 19: dCPA/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant. The size of the dots indicates the number of included flights in the scenario.

participants. While this could be an outlier, further analysis revealed that for both these scenarios this participant did not include any flights, while the other three participants included at least two. As the 7-10 NM range is just over the dCPA distance for which action is usually taken it seems that this participant decided no action was required, resulting in no included flights and a low complexity score for the scenarios. Despite this apparent outlier, the correlation was found to be weak but significant. For the Munster sector no trend is visible.

4) tLOS: Similar to the dCPA metric a negative trend can be seen for the tLOS metric and the Z-score for the Brussels sector scenarios in Figure 20. For the Brussels sector the geometry of the first interaction of the FOI seems to have a larger impact on the experienced complexity than for the Munster sector, where again this negative trend is not visible. However, even for the Brussels sector only a very weak correlation exist according to the Kendall-Tau test.

5) Required flight level change: The scenario type comparison already revealed that a required flight level change of the FOI increased the experienced complexity. The average complexity score lines in Figure 21 indicate that in the Brussels scenarios the reported complexity increased with increasing flight level change, this result is visible for both required climb and descent. The Kendall-Tau test was performed on the absolute value of the level change, and this positive correlation is found to be weak but significant for the Brussels sector. In the Munster scenarios the effect of increased flight level change on the complexity is less visible.

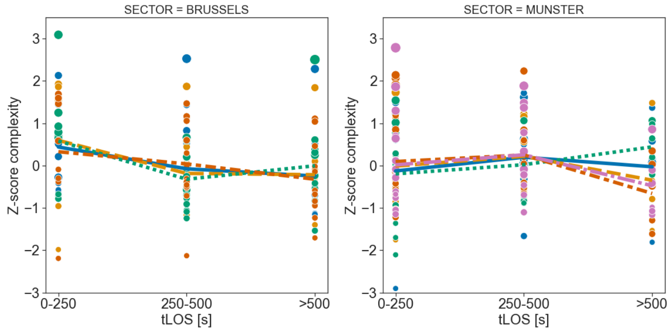


Fig. 20: tLOS/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant. The size of the dots indicates the number of included flights in the scenario.

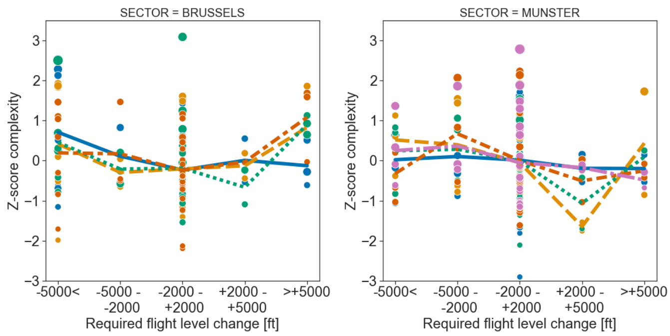


Fig. 21: Required flight level change/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant. The size of the dots indicates the number of included flights in the scenario.

TABLE IV: Kendall-Tau correlation for all scenarios per sector (significant correlations are in boldface)

sector	metric	R	p
Brussels	number of LOS situations	**0.38**	**p < 0.001**
	Number of crossings	0.01	0.91
	dCPA	**-0.16**	**p < 0.001**
	tLOS	**-0.27**	**p < 0.001**
	Required FL change	**0.27**	**p < 0.001**
Jever	number of LOS situations	0.01	0.89
	Number of crossings	0.01	0.91
	dCPA	-0.03	0.50
	tLOS	0.05	0.35
	Required FL change	**0.23**	**p < 0.001**
Munster	number of LOS situations	-0.05	0.37
	Number of crossings	**0.24**	**p < 0.001**
	dCPA	-0.04	0.44
	tLOS	0.02	0.63
	Required FL change	0.06	0.31

D. Questionnaire Response

Results from the questionnaire provide some additional insight into the reasoning behind the reported complexity scores and included flights. For each reviewed scenario the participants were asked to indicate what factors contributed to the complexity of the included flights. Unfortunately, due to technical issues part of the data from the questionnaire was lost, therefore this information is only available for 113 of 828 included flights. Many of these flights are from the Brussels sector, as the saved results were mostly from three participants of the Brussels sector, for all other participants the reviews of only one or two scenarios were saved.

There were six factors provided for the ATCo, multiple factors could be indicated per flight: Possible loss of separation between the FOI and the included flight (1), the solution space available for this flight (2), the flight level of this flight (3), he fact that this flight is climbing or descending (4), the time remaining before loss of separation (5), the XFL of this flight (6).

In Figure 22 the number of times each factor was given is plotted. In line with the results of the included flight analysis the most frequently given reason for including a flight was the XFL of a flight, and the third most given reason was the current flight level of the flight, both factors related to the possible intersection of flight levels. The second most given reason was the possibility of a loss of separation between the FOI and the included flight. One ATCo stated that he would "always give possible loss of separation as a reason, as this is the whole purpose of my job". For some scenarios the ATCo indicated that the location of the FOI or the lack of a destination airport was unexpected and not in line with the sector procedure, which added to the complexity.

Finally, the ATCos were asked to indicate how comfortable they would feel with an automated system handling the FOI. In line with the MUAC vision, the results in Figure 23 show that the ATCos would be more comfortable with automation taking control of flights that they experience as less complex. When asked to elaborate on this rating a reoccurring comment was that the level of comfort increased when the degrees of freedom for the FOI were limited. If the FOI was on a direct track to the XCOP flying at the XFL the ATCos felt more secure in their decision to let the flight be handled by the automation as the level of uncertainty of this flight was low. They argued that if the uncertainty was low a possible conflict situation could be handled by manoeuvring another flight around the FOI.

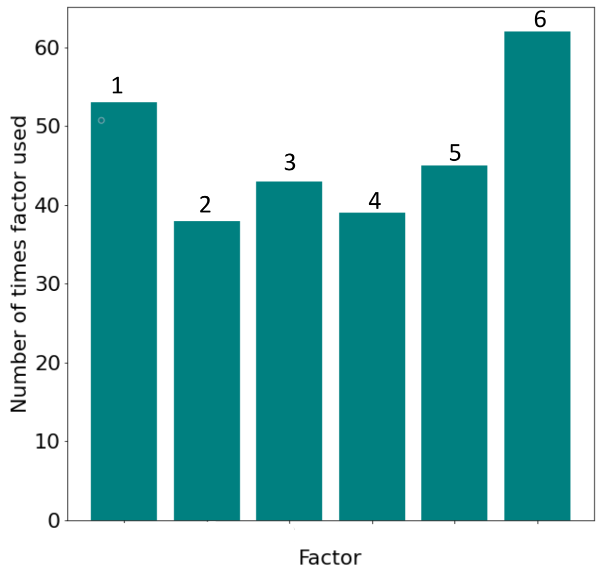


Fig. 22: Factors that contributed to the complexity of a flight.

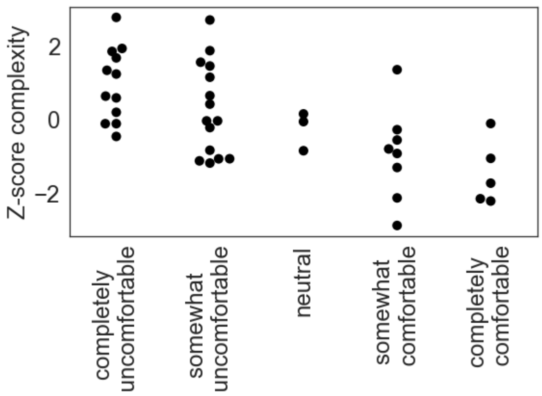


Fig. 23: Responses from the questionnaire about the level of comfort for the automation of the FOI.

VI. Discussion

A. Included Flights

A correlation was found between the number of included flights and the reported individual flight complexity for all three sectors. The difficulty with this factor is that it is not an objective measurement of the task complexity, as the included flights were an input by the ATCos. However, the included flights can give more direction in the search for a complexity model. The significance of the number of flights included shows that complexity depends on complex flights interacting with the FOI, especially with intersecting flights levels and when coming within 10 NM of the FOI in the horizontal plane. Current models for sector complexity include characteristics of all flights in the sector [7, 19, 14], however, for the individual flight complexity the system can be reduced to just the flights that interact with the FOI. The included flight analysis showed that more flights should be taken into account than initially expected at the start of this research. To illustrate, from both literature and conversations with the MUAC ATCos it was understood that checking for sufficient vertical space requires low cognitive effort, and flights with sufficient vertical separation are considered as low complexity flights. However, flights were often included when they could, at some point in the sector, be at the same flight level as the FOI. While checking for sufficient vertical space may require little cognitive effort, keeping track of two flights with the potential to lose vertical separation seems to significantly add to the experienced complexity. Following this rationale the effect of the required flight level change can also be explained, as this adds a dimension to trajectory prediction calculation performed by the ATCos and increases the level of uncertainty in these prediction, increasing the difficulty of the interaction [16, 18].

For future research it may be of interest to gain a better understanding of the criteria, for instance in terms of interaction geometry, that the ATCos used in this flight inclusion process. Especially since the number of included flights varied between participants more research is needed into which interactions most dominantly affect the complexity for a model of individual flight complexity.

B. Metric Predictive Validity

The number of LOS situations and crossing metrics did not show a correlation for all three sectors, which further confirms the finding that the system of flights considered in the complexity assessment by the ATCos is larger than expected. Therefore, Hypotheses 1 and 2 have to be rejected with the definitions of interaction and crossing used for this study. However, it should be noted that the interdependency of the metrics has to be further investigated. In the current experiment there were many differences between the scenarios that were not reflected by a single metric. The number of crossings metric, for instance, did not take into account what part of the crossings was actually a potential conflict situations, which can be an explanation for the wide range of scores found for scenarios with the same number of crossings (or any other metric for that matter).

Another result that suggests that the entire system of flights interacting with the FOI should be considered for a complexity prediction model, is the lack of a correlation between the dCPA and the tLOS metrics and the complexity. These metrics were based only on the interaction geometry for the first interaction on the current trajectory of the FOI, which appears to be less relevant to the complexity assessment, than the total number of situations the ATCOs consider. Again, while a trend was visible for these two metric for the Brussels sector Hypothesis 3 and 4 should be rejected. As the included flight analysis did show that flights with a certain proximity range around the FOI were included significantly more a metric that combines the dCPA and tLOS may show better performance than the two separate metrics.

Out of the five metrics the the required flight level change metric showed the highest correlation with the complexity. A significant correlation was found for both the Brussels and the Jever sector. The relation was not significant for the Munster sector, however, the comparison of the scenario types showed that the scenarios with a flight level change were rated more complex for all sectors. Therefore, Hypothesis 5 is accepted. Because of this difference in average scores between the scenarios types it may be of interest to analyse the performance of the metrics separately for scenarios with and without a required flight level change, again to further assess the interdependency of the metrics.

C. Automation Effect

In the questionnaire the ATCos indicated that low uncertainty for a flight reduced the complexity and increased the willingness to let the flight be handled by an automated system. An important note on this finding is that the ATCos were asked to consider a scenario where they would control all flights except the FOI. Thus, it would be easy for them to solve any potential conflict. The result of this question may be different if the ATCos were told to consider a situation where a larger part of the traffic was handled by automation, a previous study has also found that the ATCos expect interactions between human and machine controlled flights to potentially add to the uncertainty, and by extension complexity, of the traffic scenario [6].

As the level of automation used in the ATM system increases, following the SESAR vision, it will become inevitable that human and machine controlled flights operate within the same airspace. Thus, future research of complexity should include the effects on the complexity for a controller when a dealing with a partly automated traffic picture.

D. Use of Realistic Scenarios

The use of realistic scenarios affected the degree to which the independent variables could be isolated, making it more difficult to assess the significance of the individual metrics. However, this approach also revealed the impact of the combined interactions on the trajectory of the FOI, a result that may not have been found with more simplified scenarios. To counter this effect follow-up studies could be performed using scenarios that are designed such that only one isolated factor changes. Or, in case of realistic scenarios, the number of scenarios should be increased until groups of scenarios with similar characteristics can be formed and comparing between groups is possible.

The ATCos were all given scenarios from the sector for which they were licensed, and thus were intimately familiar with the structure and procedures of this sector. This did seem to add another dimension to the discussion of complexity, as ATCos mentioned on some occasions that they found a situation more complex as it was simply uncommon, and not fitting with the procedures used in the sector. Also, the FOIs were only given an XCOP and no destination airport, to constrain the ATCos to focus their assessment only on the task of bringing the FOI to the XCOP. This method in some cases seemed to add to the complexity, as it was not in line with the ATCos' way of working, which was destination based. If the ATCos had been asked to work with a fabricated sector, or even a sector for which they were not familiar with the procedures, this effect may have been mitigated.

Whether the effects of uncommon procedures have to be included in further studies of the complexity model depends, mainly, on the desired use of the model. The patterns from the current airspace sectors may not be relevant when changes to the airspace structure are made, and also a sector dependency in the model would be undesirable.

E. Between Sector Differences

The metrics performed best for the Brussels sector, where a significant correlation with complexity for all metrics except the number of crossings was found. This result is assumed to also be related to the difference in procedures and strategies used in the sector. The Brussels sector is relatively small, meaning that the look-ahead time used is smaller and possible conflict situations are solved in a more ad hoc manner. Because of this strategy the relative impact of the first interaction on the FOI's trajectory is higher, hence a more significant correlation was found between the dCPA and tLOS of this interaction and the complexity score.

VII. Conclusion

The results from the human-in-the-loop experiment showed that the complexity of an individual flight depends on the number of flights an ATCo includes in the planning of the trajectory for this flight. A required flight level change for both the considered flight and flights it interacts with adds to the experienced complexity, especially in the case of overlapping flight levels, as the uncertainty of the trajectory predictions increases. Most flights that are predicted to come within 10 NM of the individual flight were shown to affect the complexity assessment. The proposed metrics, that considered only a single traffic characteristic, were too limited for the prediction of complexity. Their sector and scenario type dependence indicated that a combination of factors is needed for the definition of complexity. Future research should focus on the correlation between the different metrics and the further understanding of which interactions most dominantly affect the complexity of an individual flight.

References

[1] F. J. Sáez Nieto, "The long journey toward a higher level of automation in ATM as safety critical, sociotechnical and multi-Agent system," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 230, no. 9, pp. 1533–1547, 2016.

[2] T. Prevot, J. R. Homola, L. H. Martin, J. S. Mercer, and C. D. Cabrall, "Toward Automated Air Traffic Control-Investigating a Fundamental Paradigm Shift in Human/Systems Interaction," *International Journal of Human-Computer Interaction*, vol. 28, no. 2, pp. 77–98, 2012.

[3] SESAR Joint Undertaking, "ATM Master Plan," Tech. Rep., 2020. [Online]. Available: https://www.sesarju.eu/masterplan2020

[4] M. R. Endsley and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," *Human Factors*, vol. 37, no. 2, pp. 381–394, 1995.

[5] EUROCONTROL Maastricht Upper Area Control Centre, "EUROCONTROL Maastricht Upper Area Control Centre OPS & Automation Strategy," Tech. Rep., 10 2019. [Online]. Available: https://skybrary.aero/bookshelf/books/5341.pdf

[6] G. de Rooij, C. Borst, M. M. van Paassen, and M. Mulder, "Flight Allocation in Shared Human-Automation En-Route Air Traffic Control." 21st International Symposium on Aviation Psychology, ISAP 2021 ; Conference date: 18-05-2021 Through 21-05-2021, 2021.

[7] P. H. Kopardekar and S. Magyarits, "Measurement and prediction of dynamic density," *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar, Budapest*, no. Dd, 2003.

[8] K. Lee, E. Feron, and A. Pritchett, "Describing Airspace Complexity: Airspace Response to Disturbances," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 1, pp. 210–222, 2009. [Online]. Available: https://doi.org/10.2514/1.36308

[9] R. Mogford, J. Guttman, S. Morrow, and P. Kopardekar, "The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature," *Storming Media*, p. 32, 9 1995.

[10] EUROCONTROL Maastricht Upper Area Control Centre, "ARGOS ORD," Tech. Rep., 1 2022.

[11] M. Prandini, L. Piroddi, S. Puechmorel, and S. L. Brázdilová, "Toward air traffic complexity assessment in new generation air traffic management systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 809 – 818, 2011.

[12] C. Boag, A. Neal, S. Loft, and G. Halford, "An analysis of relational complexity in an air traffic control conflict detection task," *Ergonomics*, vol. 49, pp. 1508–1526, 8 2006.

[13] K. A. Kimball, M. A. Hofmann, and R. O. Nossaman, "Differential velocity and time prediction of motion." *Perceptual and motor skills*, vol. 36, no. 3, pp. 935–945, 1973.

[14] B. Hilburn, "Cognitive Complexity in Air Traffic Control a Literature Review," *Eurocontrol*, vol. 12, no. 13, 2004.

[15] R. W. Remington, J. C. Johnston, E. Ruthruff, M. Gold, and M. Romera, "Visual search in complex displays: Factors affecting conflict detection by air traffic controllers," *Human Factors*, vol. 42, no. 3, pp. 349–66, 2000.

[16] C. Niessen and K. Eyferth, "A model of the air traffic controller's picture," *Safety Science*, vol. 37, no. 2-3, pp. 187–202, 3 2001.

[17] L. Pfeiffer, N. H. Muller, and P. Rosenthal, "A survey of visual and interactive methods for air traffic control data," in *Proceedings of the International Conference on Information Visualisation*, vol. 2015-September, 2015.

[18] E. M. Rantanen and A. Nunes, "Hierarchical conflict detection in air traffic control," *International Journal of Aviation Psychology*, vol. 15, no. 4, pp. 339–362, 2005.

[19] P. Kopardekar, "Dynamic Density – A Review of Proposed Variables," Tech. Rep., 2000. [Online]. Available: http://www.tc.faa.gov/acb300/Techreports/all_dds.pdf

II

Literature report (Graded for AE4020)

1

Introduction

In the past decades, all services in the Air Traffic Management (ATM) sector have managed to handle a significant growth in air traffic. From 1995 until the start of the Covid-19 pandemic in 2020, on average passenger traffic saw an annual increase of 5 percent [**?**]. Despite the current uncertainties due to the pandemic, air traffic is expected to recover to the pre-pandemic level around 2024 [19]. At this level the ATM system showed clear signs of saturation, which results in delays, rerouting and generally inefficient operations [44]. To reduce the pressure on the ATM system and reduce the environmental impact of air travel, operations need to become more efficient. Solution directions include the implementation of new technology and increased cooperation between different service providers.

The predicted increase in air traffic, growing environmental concerns and the expected introduction of various highly autonomous air vehicles, such as drones, into the airspace are all factors that contributed to the development of the Single European Sky ATM Research (SESAR) vision [46]. Currently, the airspace is divided into different sectors managed by Air Traffic Service Providers (ANSP), responsible for separation of traffic in their sector and the safe hand-off of traffic to the next sector. The SESAR vision is to create an ATM system that is dynamically scalable to demand, and decoupled from geographical sector boundaries. In the new system, with the step-wise introduction of automation and increased data sharing, ATM will be based on the principle of trajectory based operations (TBO), which allows airspace users to plan and fly their preferred 4D trajectory from gate to gate.

The introduction of automation in Air Traffic Control (ATC) will in time change the Air Traffic Controller's (ATCo) job from a manual to a supervisory control task. Research, however, has shown that humans lack the ability to adequately perform a monitoring task for a sustained period of time [15]. One common concern is the out-of-the-loop phenomenon, a lack of control loop involvement of the human operator [6]. This lack of involvement and understanding can lead to a delayed response from the human controller in case of an automation failure. To mitigate the negative effects of automation, adaptive automation principles will be introduced that can dynamically guide the division of tasks between the system and the controller. More specifically, routine and repetitive tasks will be handled by the automation system, while the ATCo performs the more complex work. The objective for this adaptive automation is to maximise human performance and engagement, and reduce workload [46].

Maastricht Upper Area Control Centre (MUAC), an Area Control Centre (ACC) that manages traffic above 24,500 ft over the Netherlands, Belgium, Luxembourg and part of Germany, is exploring the introduction of more automation into the ATC workdomain to contribute to realising the SESAR vision. In line with the SESAR plan MUAC's automation strategy has several layers, with different levels of automation on the way to fully automated ATC. MUAC is developing a system that should, in the future, autonomously handle traffic. The system, named ARGOS, will initially only control "basic" flights, while the human controller handles the remaining, more complex, traffic. The reasoning behind this division is twofold. First, the ATCo is kept engaged and in the loop by performing the more cognitively demanding tasks. And second, the automation system will handle "basic" flights for which strategies of safe-guarding the separation minima are well within its capabilities, increasing the overall reliability of the system. For this approach to work, an algorithm is needed that determines which flights are "basic", and can safely be handled by ARGOS, and which flights are "non-basic".

An exploratory study by de Rooij et al. showed that a mixed responsibility scenario, where part of the traffic is handled by the ATCo and the other part by the automation system, is feasible. However, this study recommended further research into factors that can be used to determine which flights are "basic" or "non-basic" [13]. Past research on the complexity of air traffic has been focused on the measurement of the combined complexity of all traffic in a sector [25], [26], [28]. During this research project the previously identified factors for sector complexity will be translated to factors that contribute to complexity for an individual flight. For this thesis this information is used to develop an algorithm, that assigns a complexity score to each individual flight, which can be used label a flight as either "basic" or "non-basic".

1.1. Research objective and scope

The research objective for this project is:

To find factors that can be used to predict the complexity of an individual en-route flight by designing a flight complexity metric, with the intention of distributing flights between manual control and automation, and validating it with simulator experiment.

This research will be limited to the study of en-route traffic, meaning complexity related to the traffic moving from or towards an airport will not be included. For the experiment a TU Delft-built Java-based simulator will be used. The flight complexity metric determines the complexity by making calculations based on the aircraft trajectory, in the simulation the position of each flight can be accurately predicted throughout the duration of the scenario. In practice not all information used is currently available to the ANSPs, however, the planned 4D trajectory based operations and increasing automation capabilities should provide the needed information to make these predictions in the future.

1.2. Research question

The main research question is for this project is:

What factors affect the perceived complexity of an individual en-route flight by an ATCo and how can these factors be used in an algorithm to define a complexity score that can be used to distribute flights between manual control and automation?

A number of sub-questions were drafted to provide structure to the research project. The first question guides the literature study. The third and fourth questions are related to the design and validation of the flight complexity metric.

1. How is air traffic complexity currently defined?
    (a) How is air traffic complexity defined and how is it related to controller workload?
    (b) How do different ATCo tasks affect perceived complexity?
    (c) To what extent is perceived complexity affected by individual differences between ATCos?
    (d) What methods are currently used to define air traffic complexity?
2. What traffic characteristics are most relevant to determine the perceived complexity of an individual flight in a traffic scenario?
    (a) How can methods used to find sector complexity be translated to find the complexity of an individual flight?
    (b) What traffic interaction properties affect single flight perceived complexity?
    (c) What sector properties affect single flight perceived complexity?
    (d) How does the solution space of a flight impact the single flight perceived complexity?

3. How can a complexity metric adequately identify a flight in a traffic scenario as basic or non-basic?
   (a) What flight (track) information is needed to determine flight complexity?
   (b) How can complexity of a flight be mathematically represented?
   (c) How can a threshold be defined to label flights as basic or non-basic?

1.3. Report outline

The first part of the thesis project is the literature study. The literature study has main four topics focused identification of a definition for flight complexity. The first topic, discussed in Chapter 2, is the use of automation in a human-in-the-loop system is. More specifically, the use of automation in the ATM sector. This chapter explains the effects of introducing automation and the feasibility of a shared responsibility traffic scenario. Chapter 3 is an analysis the ATCo workdomain. In this chapter the relation between different ATCo processes and traffic complexity is assessed. Chapter 4 discusses current models used to define the ATCo workload. In these models the complexity of the air traffic is one of the factors that affect the perceived workload. The perceived workload is also subject to individual differences between controller and can be affect by the set-up of the ATCo workdomain and the ATC procedures. Furthermore, this chapter discusses different metrics that are currently used to measure the complexity of air traffic. The metrics mentioned have been reviewed for their performance as a metric for sector complexity. However, some the factors included can also be used to define flight complexity. The last topic of the literature study is ARGOS. Chapter **??** focuses on the design objectives and logic of ARGOS. This logic shows the capabilities and preferred resolutions of ARGOS.

In Chapter 5 three potential metrics are introduced to define flight the complexity of single flight. An experiment is designed to find the correlation between these metrics and the subjective ATCo score of of complexity. The experiment design is discussed in Chapter 6.

2
Automation in human-in-the-loop control systems

Research into the topic of automation and, more specifically, human-automation interaction, has been focused on several distinct themes in the past decades. An important theme throughout is managing the detrimental effects that can occur with the introduction of automation into the control loop. Thus, for the design of an automated system it is important to be familiar with these potential threats to the performance of the human controller. This chapter discusses some key lessons from research from the introduction of automation in generic human-machine systems. Also, concentrating on the topic of this thesis project the automation strategy for ATM and the automation tools currently implemented in the ATCo workdomain are discussed.

Section 2.1 is an introduction of automation research. Next, Section 2.2 explains the SESAR vision for increasing automation in the ATC sector. Sections 2.3 and 2.4 discuss the negative effects of automation, and several mitigation strategies. Finally, Section 2.5 discusses findings from simulations with different shared responsibility concepts.

2.1. Automation research

Automation is defined by the Britannica as the: *"application of machines to tasks once performed by human beings or, increasingly, to tasks that would otherwise be impossible. Although the term mechanisation is often used to refer to the simple replacement of human labour by machines, automation generally implies the integration of machines into a self-governing system"* [27]. While this definition makes no mention of a computer, it is the advancement in computer processors that has driven the introduction of automation.

The uses of automation have greatly expanded over the years, which is reflected in the most common topics in automation research in the past decades. Starting in the 1970s and 1980s with research focused on the automation of knowledge acquisition, the gathering of knowledge from an expert to be used in a computer programme. This evolved in the 1990s to the study of automation to aid humans in time-sensitive decision making. More recent studies investigate automation in situated embodied agents, and the introduction of automation into the lives of non-professional users. Janssen et al. performed an analysis of the human-automation interaction research field to identify current and future themes in automation research [22]. They found three themes that have been discussed extensively in previous literature and will, according to Janssen et al., stay relevant with the further introduction of automation, due to the introduction of automation in safety-critical areas and the increased availability of automation to non-professional users. The first theme, which is also important to this research project, is function allocation of tasks between human and automation. A second, related theme, is the focus, attention division and attention management of the user. Again, topics that are of great concern for the introduction of automation in ATM. The last theme that is covered extensively in automation research is how to promote the correct levels of trust, avoid misuse and avoid confusion of the user [22].

2.2. The SESAR vision for automation in ATC

As a significant increase in air traffic is expected in the coming decades both the European and the American aviation authorities have introduced plans to use automation in the ATC system to increase the airspace capacity [46]. In European airspace steps are taken to introduce automation in all aspects of the ATM system according to the SESAR vision. The final goals of SESAR is to create a seamless and safe environment, where both manned and unmanned aerial vehicles can operate. The same airspace infrastructure and services will be available for all airspace users. As a part of this vision the operation of air traffic will become more flight-centric. This TBO should make it possible for airspace users to plan their preferred optimised trajectory from gate to gate, reducing fuel consumption. Currently, flights are being directed by different operators in different phases of flight. TBO will also increase the predictability of the air traffic, enabling the possibility of reduced buffers, which increases airspace capacity. Furthermore, with the knowledge of the all traffic trajectories the airspace can be more dynamically adjusted for the expected demand. This entire system hinges on the system wide sharing of information and the development of new communication tools.

Figure 2.1 shows an overview from the SESAR Master Plan (MP), for the gradual introduction of automation in four system function: information acquisition and exchange, information analysis, decision and action selection, and action implementation. The SESAR MP is split into four phases, the last columns give an overview of level of automation introduced in each phase. The last column mentions U-Space, this is a new branch in the ATM system that will be introduced to manage unmanned aerial vehicles. In the first phases the automation merely has a support function and cannot initiate any action. With increasing automation the system becomes more autonomous, until full autonomy is reached in level five.

The productivity of air navigation service providers is expected to increase due to the introduction of automation into ATC. The easier and repetitive tasks of the ATCos will be automated leaving them to perform the more complex work. Their work will be supported with more separation management tools made possible by the increased information sharing and trajectory information [46].

2.2.1. Current automation in en-route ATC

The ATCos at MUAC, currently, have some automation support tools available. Each of these tools aids in the detection of an interaction between flights, or with restricted airspace or terrain. These tools have been introduced or updated in the first phases of the SESAR MP. They are the lowest level of automation, as they support the ATCo only in the information acquisition and analysis, by analysing the traffic scenario and finding potential threats to the safety of a flight. The tools differ in the look-ahead time and information sources used:

Short Term Conflict Alert (STCA): The STCA tool is a safety net tool that warns the ATCo for an imminent loss of separation, the look-ahead time for the STCA is 1-2 minutes. The input for the STCA is the surveillance data and the tool uses the assumption that the aircraft will remain on the current track. Some versions of the STCA include flight level data from the SSR mode S or ATCo input [49].

Tactical Controller Tool (TCT): The TCT warns the ATCo in case of a potential conflict in the sectors. The inputs for this tool are the current aircraft track and the current aircraft trajectory. The TCT gives potential conflict warnings in both the horizontal and the vertical plane. This tool also aims to reduce workload by indicating when there are no pending conflicts in the sector. As a secondary function the TCT can give warnings for critical missed manoeuvres, a situation where a loss of separation would occur if a pilot would not execute a given instruction. The look-ahead time of the TCT is 5-8 minutes [49].

Medium Term Conflict Detection (MTCD): The MTCD has a look-ahead time for 20-30 minutes and can make use of any information system to make predictions. This tool notifies the ATCo in case of a possible loss of separation between two aircraft, and when an aircraft potentially enters a restricted airspace. The conflict detection also includes aircraft that do not violate the separation minima on their current heading and flight level, but that are at risk to be in a conflict with a pilot requested clearance or if a clearance to solve a different conflict is given [49].

Verification and Resolution Advisory (VERA): The VERA tool extrapolates the aircraft positions based on surveillance data and assists the ATCo to determine whether two aircraft will be in conflict if they were to continue on their current headings [9].

Area Proximity Warning (APW): The APW tools gives a warning when an aircraft is predicted to fly into a notified airspace, which is a restricted area. The APW system uses surveillance data and flight path prediction. It is a short term alert tool with a look-ahead time of up to 2 minutes [48].

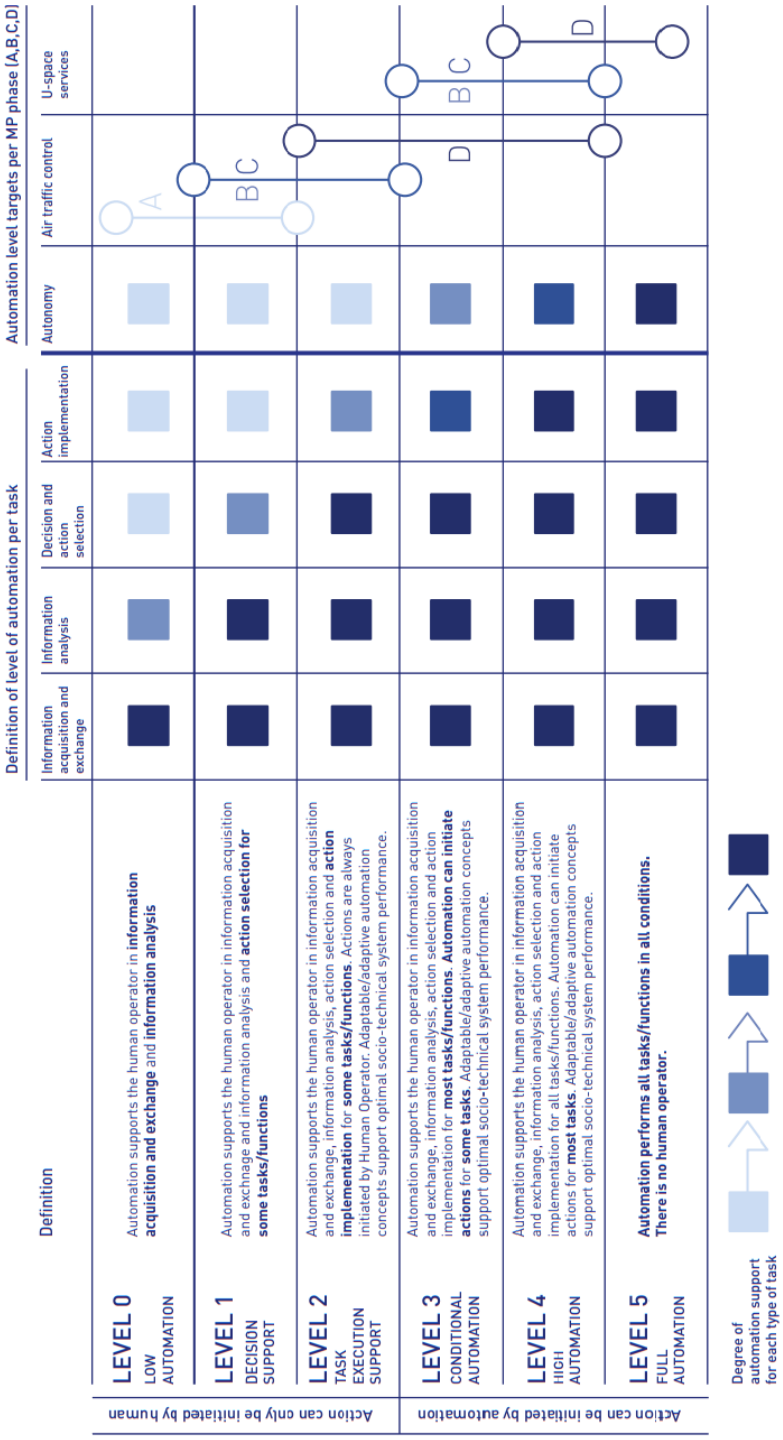


Figure 2.1: Levels of automation in SESAR Master Plan [46]

2.3. Automation challenges

Despite the potential benefits, past research has taught us that automation does come with several (unexpected) difficulties. Bainbridge [5] described some of these problems or, in her words, ironies of automation back in 1983 and many of them are still relevant to this date [54]. She states that in the design of an automated system the engineer plans to make the human operator obsolete, however, as a result the operator is left with the tasks that the engineer was unable to automate. Instead of eliminating the operator, the system is now dependent on the controller to handle only these tasks, that where deemed to difficult to automate. And even a fully automated system is no guarantee for reliability, as errors in the design process will lead to errors in the operating of the system, which then have to be solved by the very operator that was deemed unreliable.

With increasing automation the role of the ATCo will evolve from a manual to a supervisory control task. Research shows that humans have difficulty to adequately perform a supervisory task for a sustained period of time [15]. Currently, one of the main concerns when it comes to the introduction of automation is the Out-Of-The-Loop (OOTL) phenomenon, which is a lack of control loop involvement of the human operator [6]. In a highly automated control environment the automation performs most of the manual tasks that were previously assigned to the human controller. As the human supervises the process from a high level, a certain distance occurs between the human and the control loop. This distance from the control loop adversely affects the response of the controller in an abnormal situation where he/she needs to take back manual control.

One of the leading causes for OOTL behaviour is the loss of situation awareness. Situation awareness is defined by Endsley as *"the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* [16]. According to Endsley, situation awareness can be split into three levels, the first level is the perception of the elements in the system, the second is comprehension of the meaning of the system with the elements perceived, and the third level is being able to make a prediction for the elements in the system. All three levels are relevant to the timely conflict detection and resolution task of the ATCo and the occurrence of OOTL behaviour needs to be avoided when designing new automation tools.

Several factors have been found can that contribute to the loss of situational awareness:

Vigilance: In a system with a high level of reliable automation the human task is to monitor the system for rare anomalies, and to spot these anomalies a high level of vigilance is required. However, studies on attention show that when the operator is performing a monitoring task, vigilance decreases over time [15]. While vigilance problems are sometimes assumed to be related just to the boredom of performing a simple task, Parasuraman showed that also for monitoring complex tasks the same vigilance effects can be found [31].

Complacency: Complacency occurs when the operator relies too much on the automation performance. Especially before the first system failure human operators are shown to have a tendency to overestimate the reliability of the system. The effects of complacency are particularly present when automation has a constant reliability, while a system with varying reliability has been shown to keep the human operator alert to potential automation errors [33].

Passive information processing: In a manual control task the operator is actively processing the information from the system feedback, while for a supervisory control task the operator passively processes the feedback. This passive processing leads to a latency in the response when an error in the system is detected, as in the case of passive processing the required information is not readily available from the active memory [50].

Despite being trained for the manual control task, operator performance declines without the opportunity to practise the task themselves [59]. This phenomenon is known as skill erosion. When an operator is trained with a control system that already includes a high level of automation, he/she may not even lack the needed control skills altogether. This issue is relevant to ATCos today as they are trained with several automated support tools [54]. However, training the controller to be better adapted to the human-machine interaction can also make him/her better prepared for working with a highly automated system and decrease the occurrence of the OOTL behaviour [4].

2.4. Mitigation of automation issues

To design a properly functioning automated system both the technical and human capabilities have to be considered. With the right level of automation the negative effects of automation can be mitigated. Parasuraman et al. [34], therefore, propose a model to determine what part of a control system can be automated and to what extent. The level of automation in this model is ranked on a ten-point scale, where level one is no automation at all and level ten is a fully automated system. Figure 2.2 shows the definition of the different automation levels. The presented definitions mostly refer to the automation of the decision making process. As a control system includes more processes, four stages were defined based on human information processing, to extend the model.

These stages can be translated to the following system functions:

1. **Sensory processing:** The first stage is the sensory processing, which refers to the acquisition of information from the environment.
2. **Perception/Working memory:** The second stage is the perception of the information, or information analysis as a system function. Within this stage possible solutions can be generated, but no decision is made.
3. **Decision making:** The third stage is decision making and determining the required action to perform.
4. **Response selection:** The final stage is the implementation of the selected action.

The level of automation can be different for each of the stages in the model, and Parasuraman et al. present two different sets of criteria to determine to what extent level a system function should be automated.

The first set of criteria describes the effects of automation on human performance. The effects of the automation on the mental workload, but also on situation awareness, complacency and skill degradation need to be considered.

The second set of evaluation criteria combines the reliability of the automation and the potential cost, or risk of a decision or action taken by the system. Choosing the wrong level of automation for the reliability of the system affects the trust of the human operator in the system, which can negatively impact the use of the automation. In case the operator mistrusts the system the automation might get disabled, conversely, when the operator over-trusts the system he/she might not respond properly to system errors. The cost of the decision refers to the impact of an action taken by the automation. High levels of automation are appropriate for tasks with low impact, where the cost of a wrong action is low. Another implementation option of high level automation is in an emergency situation where an immediate response is required and there is insufficient time to wait for a human operator to make a decision; these decisions can be higher impact.

Apart from designing a system with the correct level of automation, several other policies can be used in the mitigation of automation issues:

Human operator adaptation: Training the human operator to perform better with automation and be more aware of the risks of the OOTL phenomenon can decrease controller error. Bahner et al. showed that when controllers are introduced during training to rare system errors, complacency reduces [4]. Bruder et al. suggest that as more automation is introduced in the work of an ATCo, the selection criteria for this job will have to evolve. The goal of their study is to create a metric to select applicants that perform better in an automated environment [10].

System adaptation: System adaptation is the practice of designing an automated system to work as a partner to the human operator, and not simply let the system perform tasks that were previously done by the human. Automating the controller's manual tasks can lead to the loss of feedback from the system, and consequently the loss of understanding of the system by the operator [17]. Cook et al. state that as the system communicates with the human through displays, it is extremely important that relevant information is provided to the human [11]. Thus, new functions should be designed that support the human operator. For successful human-machine cooperation the status and objective of the automation functions should be made clear to the controller, and the automation should be designed to guide the allocation of attention of the controller [6].

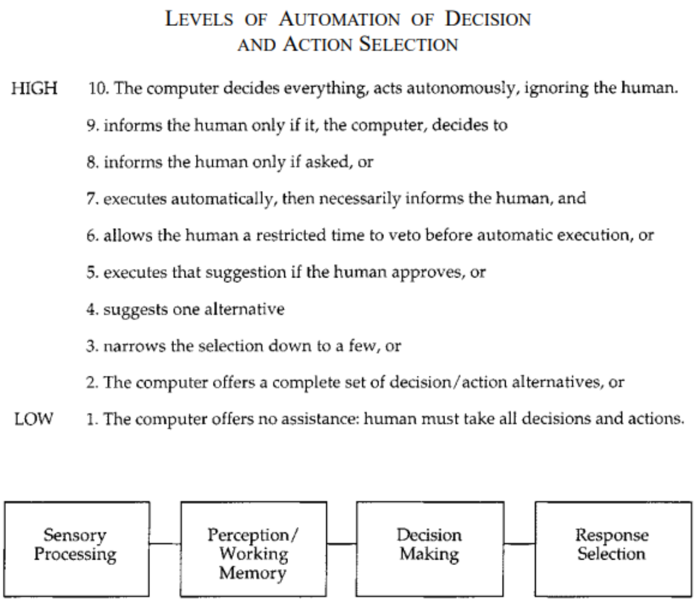


Figure 2.2: Levels of automation and stages of the control process [34]

Adaptive or adaptable automation: With adaptive automation the automation system is adapted to the need of the operators. The responsibility for this adaptation can be with the system (adaptive automation) or with the controller (adaptable automation). The level of automation and the allocation of tasks changes in real-time, taking into account the current current state of the controller. Changes can also be made in the feedback and information that is presented to the controller. This form of automation has been shown to be able to moderate controller workload [53]. In adaptive automation the system can change the function allocation depending on operator performance, the occurrence of critical events or changes in the the psycho-physiological measures [32]. In adaptable automation the operator actively controls the level of automation [45].

2.5. Shared traffic responsibility concepts

For the introduction of automation into ATC, different concepts have been developed. In most of these concepts the automation level is increased gradually, which means that on the way to full automation there are (several) stages where humans and computers share the responsibility of controlling the traffic. This section discusses findings from research into these concepts on the effects found on performance and the human operator.

NASA has experimented with the automation of separation assurance and increased levels of traffic load. This research showed that with increased aid from automation, higher levels of air traffic could be safely handled [39]. In this concept, the same level of automation support is available for all flights in the scenario, but different ATC tasks could be automated. Participants of the experiments were also asked what tasks, according to them, could be allocated by the automation system. The results of this survey are shown in Figure 2.3. A large majority of participants states that the hand-off of flights and transfer of communication can be allocated to automation. For many of the decision making, tasks the participants indicated that the responsibility could be shared between automation and controller, implying a willingness to accept the increased automation aid into the ATC workdomain.

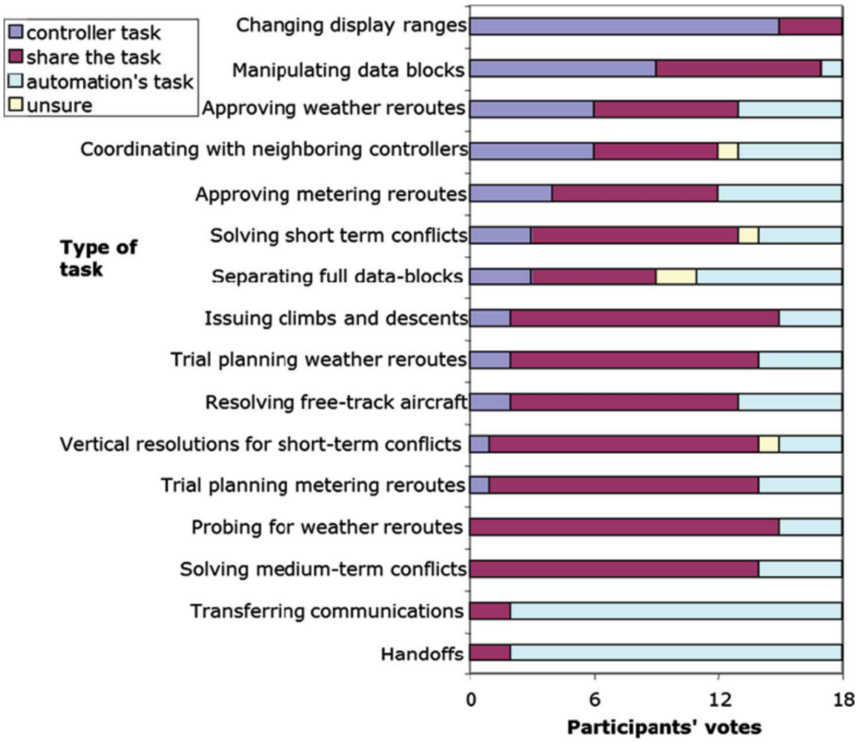


Figure 2.3: Allocation of tasks according to ATCos [39]

Different from the NASA concept, SESAR and MUAC envision to increase automation in ATC by allocating part of the traffic to automation. The automation system thus is responsible for all ATC functions of the allocated traffic. In this concept, flights that are handled by automation and flights that are managed by the ATCo operate in the airspace, which can lead to additional complex situation when automated flights interact with ATCo controlled flights [43]. Rovira et al. performed an experiment where part of the traffic was allocated to automation. Two different scenarios with different levels of intent information available were compared. In the first the ATCo had no intent information for the flights allocated to automation, in the second scenario the ATCo knew the intent of the automation system for these flights. This research showed that while the mixed traffic scenarios supported earlier detection of conflicts, the ATCos also allocated a significant amount of visual attention to the automated no intent aircraft, for which they were not responsible [43]. This indicates that for mixed responsibility scenarios the predictability of the automated traffic needs to be considered to ensure that ATCos appropriately allocate their attention. A different study with a mixed responsibility scenario also showed that ATCos experienced a reduced level of situation awareness for flights that they were not responsible for [58].

A comparison can be made between these mixed responsibility scenarios and the sectorless operation concept, as proposed by Birkmeier et al. [7]. Studies into sectorless ATM can thus provide further insight into the effects of mixed responsibility concepts. In sectorless ATM controllers are assigned to control a number of aircraft that are not geographically bound to a sector. Instead, these aircraft can be spread over a larger distance, and an ATCo is responsible for the same flight during the entire cruise phase, as is the case in the concept of TBO proposed by SESAR. In the sectorless ATM concept the flights controlled by an ATCo interact with flights that the ATCo has no control over, similar to situation presented in the mixed responsibility allocation concepts. For these situations new concepts have to be developed to determine who (or what in case of automation) is responsible for separation in case of a conflict. Simulations by DLR showed a change in the mental picture of the ATCo for the sectorless ATM concept. The mental picture is mental representation of traffic that supports the situation awareness of the ATCo. The simulations showed that for sectorless ATM the ATCo was capable of managing several mental pictures that were smaller and had a shorter timeline than the mental picture used for the current ATM concept. Figure 2.4 illustrates the shift in mental picture.

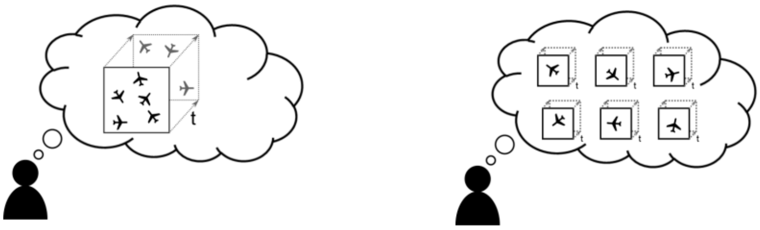


Figure 2.4: Shift in mental picture from the current ATM system (left) to the sectorless ATM concept (right) [7]

2.5.1. Allocation strategy factors

The type of traffic allocated to automation probably affects the feasibility of the mixed responsibility concept. De Rooij et al. concluded that for a mixed responsibility traffic scenario it is is important to find an allocation scheme that determines whether a flight is "basic" or "non-basic" [13]. As part of their exploratory research experiment ATCos were asked to fill out a questionnaire to indicate which factors, other than complexity, they considered most important for they allocation of flights to automation. The results showed that the responsible entity (human or machine) for the traffic directly around an aircraft is considered the most important factor. The ATCos argued that the single automated flight interacting with a number of manually handled flights would add unwanted uncertainty to the traffic situation. Another important factor found is the capability of the automation system. In this simulation the automation was not capable of issuing a direct clearance to the sector exit point instruction, the ATCos indicated that had this been the case they would have delegated more flights to the automation [13]. The results from the questionnaire can be found in Figure 2.5.

This thesis project solely focuses on the delegation of traffic to automation based on the traffic complexity. However, research shows that other factors are relevant to consider in further development stages of SESAR MP, as they have a direct impact on whether the allocation scheme is accepted by the ATCos.

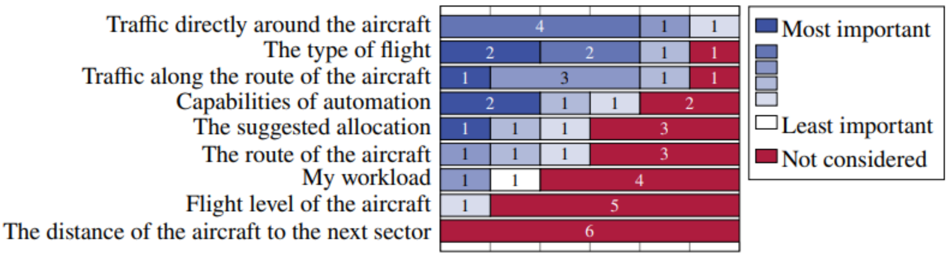


Figure 2.5: Overview of relative importance of factors for the delegation of flights to automation as indicated by ATcos [13]

3

The ATC workdomain

Aid provided by automated support tools, such as Argos, has to be well-aligned with the ATCo work processes to be accepted by the ATCo. Research regarding the design of such tools often mentions the need to understand the working principles of the ATCo. Therefore, as a basis for this thesis project this chapter discusses the current ATC workdomain.

Sections 3.1 and 3.2 discuss the ATCo workdomain and the ATCo work. Section 3.3 describes a model of the cognitive processes of the ATCo when managing the traffic. Section 3.4 further elaborates how traffic characteristics influence the conflict detection and conflict resolution task. Section 3.5 reports the key finding from this chapter and their relevance to this thesis project.

3.1. The ATCo dyad

In the current ATC system, two ATCos are responsible for the management of aircraft in a sector. Pfeiffer et al. [37] have described the division of tasks and interaction within the ATCo dyad at the Deutsche Flugsicherung (DFS). The first ATCo, the executive controller (EC), is responsible for separation of the flights in the sector and all communication with the pilots. Communication is done via radio link or Controller Pilot Data Link Communication (CPDLC) if the aircraft is equipped with this technology. The second ATCo, the coordinator controller (CC), coordinates the flights leaving or entering the sector, he/she communicates with the CC of the previous or next sector to arrange the handover of a flight. As a secondary task, the CC also verifies the communication between the EC and the pilot. The ATCos work side by side and use a radar screen, that shows a 2D representation of the traffic in their sector. The aircraft that will cross through the area of responsibility of the ATCos are shown with a flight label. Other flights in the air space are visible in a discrete light grey. The flight label has the following information:

- Aircraft ID
- Actual Flight Level: flight level that the flight is currently flying at
- An arrow to indicate if the flight is climbing or descending
- Cleared Flight Level: flight level that the flight has been cleared for
- Information on the direction of the flight, either an indication of the heading or the next waypoint on the route
- Transfer Flight Level: flight level that the flight has to reach before being transferred to the next sector

Upon entry of the sector a flight has to be assumed by the ATCos, after this happens they are responsible for the management is this flight, and can issue clearances. The assumed flights are shown in green on the radar screen, while the unassumed flights are grey. Figure 3 shows ATCo working position at MUAC [3].

Figure 3.1: Control room at MUAC [3]

3.2. ATCo Task Analysis

Kallus et al. [23] have performed a tasks analysis of the ATCo job. For the en-route task six main task processes were identified:

A **Taking over position/building up mental picture:** At the start of every shift the ATCo has to take over the sector responsibility from the leaving ATCo, Before the take-over happens the ATCo familiarises him/herself with the sector conditions, such as the weather and traffic density. Next he/she checks the radar screen and receives a briefing from the leaving ATCo. If there are no pending conflicts in the sector the ATCo can can set-up a sector plan, if there are conflicts the new ATCo will first follow the sector plan of the leaving ATCo until the conflicts are resolved, before establishing his/her own plan.

B **Monitoring:** The monitoring task involves updating the mental picture and checking the sector for conflicts. Additionally, the ATCo keeps track of the action hierarchy, to determine the most urgent conflict or requests to solve.

C **Managing routine traffic:** Routine traffic, free of conflict, is issued clearances for heading or level changes to comply with the predetermined route.

D **Managing request/assisting pilots:** The ATCo can also receive requests from the pilot that were not in the flight plan. New clearances may be given after such a request.

E **Solving conflicts:** If a potential conflict is found, the ATCo proceeds with the solving conflicts process. The ATCo does not always solve the conflict immediately after detection, waiting to solve a conflict can lead to a better prediction and solving the conflict later might lead to a more efficient solution.

F **Switching attention:** After giving a clearance the ATCo updates his/her mental picture and checks for changes in the traffic scenario, before returning to any of the previous tasks.

Within the main processes one or more of the following sub-processes are performed:

1. Confirming/updating mental picture
2. Checking
3. Searching conflicts/checking safety
4. Issuing instructions

After the initial take-over of the sector, the ATCo is constantly switching between tasks B to F. Figure 3.2 shows the relation between the tasks.

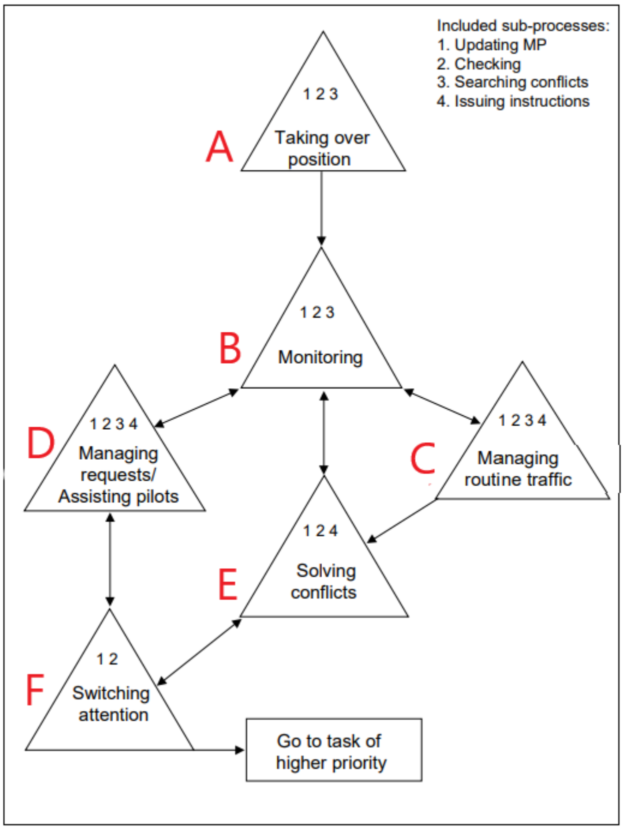


Figure 3.2: Relation between task processes of the ATC task [23]

3.3. The mental picture

One of the sub-processes from the task analysis is updating the mental picture. The "mental picture" is a term often referred to in research on the cognitive processes of an ATCo, however, how the mental picture is constructed and used by the ATCo seems to be subject to personal differences between ATCos. The following paragraph introduces the concept of the mental picture. ATCos are trained to perform multiple tasks, and deal with many sources of information all while working under time pressure [41]. The mental picture is used to structure the available information sources. By interviewing ATCos from the Deutsche Flugsicherung, Pfeiffer et al. found that the ATCos build up a picture in their head of the traffic situation in the sector by continuously scanning the position of all flights on the radars screen [36]. For this mental image, the information of the 2D radar screen is combined with information on altitude, velocity and direction from the flight label. Some literature defines this picture as a 3D model of the traffic situation [47], in other research it is noted that the dimension of the picture differs per ATCo, it has been described as 2D, 3D or even 4D, adding the dimension of time [36], [55]. ATCos need the mental image to make future predictions of the aircraft trajectories and (solutions for) possible conflicts.

3.3.1. Using the mental picture

The task analysis does not describe how the mental picture is constructed and used for conflict detection and resolution. Niessen et al. [30] have attempted to create a cognitive model that would describe the building and use of the mental picture. Initially they set out to base this model on existing theoretical frameworks for such a control task. They found that no fitting framework existed yet, as the model should account for a dynamic task environment, where states can change without the ATCo input. Also, the mental picture is created specifically to aid the ATCo in the flight separation goal, this means that the mental picture is not necessarily an exact replica of the radar screen but rather a representation of the objects is in the systems that demand the controller's attention. The model they created can be seen as an extensions of the task analysis, providing an insight in the information management of the ATCo. Figure 3.3 shows the cognitive model as presented by Niessen et al. named the Model der Fluglotsen Leistungen [30]. The model is built with six modules that surround the mental picture. The six modules are:

1. Data selection
2. Anticipation
3. Conflict resolution
4. Update
5. Sector knowledge
6. Control

All modules are connected with three information processing cycles: the monitoring cycle, anticipation cycle and conflict resolution cycle. The mental picture is a representation of the flights and the (predicted) interaction between flights in time. The mental picture is described as a working memory of the ATCo, where the different objects in the mental picture are not all equally available, but rather are activated from the long-term memory when they are the focus of the attention of the ATCo [12]. The model is further explained by the role of the three cycles:

Monitoring cycle: The data selection and update module together form the monitoring cycle, in this cycle data is selected from all available resources and used to update the mental picture. In this cycle flights can are more or less activated depending on their status. A flight can be activated for different reasons such as proximity to other flights or vertical movements, information in activated flights is updated more frequently.

Anticipation cycle: The anticipation module also forms the anticipation cycle. In this cycle the ATCo attempts to find the relation between aircraft pairs and makes a prediction regarding the future state of the flights. If a conflict is detected the aircraft pair get a mental timestamp that is indicates of the time remaining to solve the conflict. The timestamp is part of the information in the mental picture.

Conflict resolution cycle: The conflict resolution cycle sorts the identified conflicts on urgency. In order of urgency the ATCo finds an resolution, for this potential resolution the future states of the aircraft involved are predicted by "feeding" the states into the anticipation cycle. Due to the high workload it is important to solve conflicts efficiently and in a timely manner. Therefore, the ATCo first attempts to find a solution from their experience-based episodic memory, which can be seen as a list of potential solutions that are used regularly. If no suitable solution if found this way, they resort to solving the conflict using knowledge based problem solving skills [23]. If the proposed solution is deemed viable the ATCo can take action and execute the needed commands.

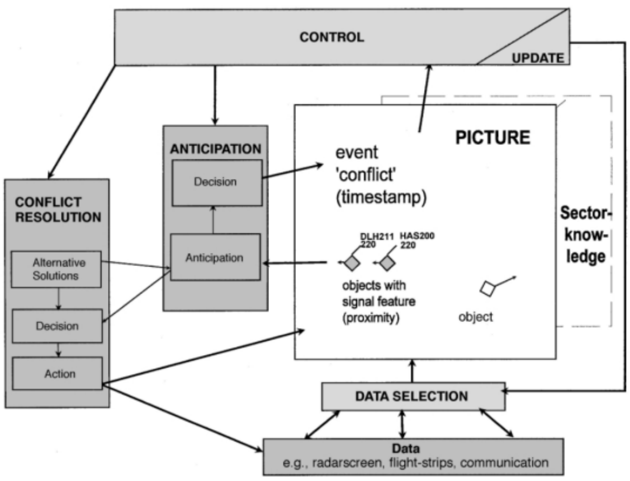


Figure 3.3: Mental model of the ATCo by Niessen et al. [30]

3.4. Conflict detection and resolution

The cognitive processes of the ATCo during conflict detection and resolution are further examined. Apart from adhering to strict separation requirements, the ATCos have some freedom in the tactics they use to manage the traffic. This freedom, in combination with performing their job under constant time pressure, while having limited mental capacity, leads to the use of different strategies and preferences to regulate the experienced workload [52]. Rantanen et al. performed an experiment to validate a hierarchical model for conflict detection which supports the premise that ATCos selectively use information to reduce the required cognitive work. This hierarchical conflict detection model describes the use of altitude, heading and speed information [41].

Altitude For any traffic situations, ATCos first check if there is sufficient vertical separation between two flights, doing this requires little cognitive effort. Rantanen et al. found that the response time for determining if two flights were in a potential conflict was the same for flights with converging and diverging headings when there was sufficient vertical separation, leading to the conclusion that checking for vertical separation is the first step in the conflict detection process.

Heading When two flights are on the same altitude, several factors influence the detection difficulty of a possible conflict. In terms of heading the flights can be on diverging, parallel or converging paths. For flights on diverging paths the separation distance increase, thus there is no conflict detection. For parallel tracks in opposite direction, the time to the closest point of approach influences the detection difficulty, as a longer time to the closest point of approach means that the flights have a large distance between them on the radar screen. For parallel paths in the same direction, the relative velocity determines whether a conflict can occur. Remington et al. found in the case of converging flight paths, larger convergence angles lead to slower conflict detection. They argue that two flights with a small convergence angle have smaller screen distance which leads to faster and more accurate detection of a conflict than flights with a large convergence angle [42]. However, other research showed that conflicts between two flights with a shallow convergence angle are more difficult to detect. The least cognitive effort is required for the detection of conflicts with a convergence angle of 90 degrees [21].

Speed While Rantanen et al. did not find a significant effect of aircraft speed on the detection time of a conflict in their experiment, integrating speed into the mental model does have a higher cognitive cost [52]. Bisseret also found that speed information was requested last en less frequently by ATCos in an experiment where information was only made available to the ATCo at their request [8]. For two flights, on converging headings travelling at different velocities, the cognitive work using the two velocities to make prediction on a conflict is higher and this leads to less accurate conflict detection [24].

Like conflict detection, conflict resolution strategies can differ due to individual differences between ATCos. These differences in resolutions strategies make it difficult to create a model of ATCo conflict resolutions. Research into conflict resolution even found differences the strategies of ATCos depending on mental state.

Despite these differences, after a conflict is detected each ATCo follows the same general path to resolve it. The first step after the detection of a conflict is to determine which of the two aircraft involved should get a resolution, it is uncommon to give a resolve a conflict by issuing instructions to both aircraft involved in the interaction. This first decision is already largely dependent on personal differences. Eyferth et al. showed the same conflicts to a group of ATCos, and asked which aircraft they would give the resolution to, the outcome of experiment yielded no basis for a model that would that could predict what aircraft the resolution would be given to [18]. However, Eyfreth et al. do state that a resolution provided by an automated support tool will only be accepted if the ATCo finds it to be sufficiently justified.

A conflict can be resolved by making a change to the heading, altitude or velocity. Changing the course of the aircraft results in a visible change to the trajectory on the radar screen. When an aircraft has to deviate from its route, at least one more instruction is needed to return the aircraft to the route, adding to the total workload. An exception here is changing the course to a direct to the exit point, after which no extra instruction is needed. If the deviation leads to a change in exit point, or change in exit time, extra coordination with ATCos from the adjacent sector is needed. The effect of a resolution on the travel scenario will be different for each individual case. Spaeth et al. conducted an experiment to find the factors that are mostly considered in choice of resolution method [51].

The factors are listed here:

- Destination
- Aircraft performance
- Vertical distance to sector boundary
- Lateral distance to sector boundary
- Attention stress
- Coordination effort
- Sector load
- Weather
- Procedures with adjacent sector
- Distance between conflicting aircraft

The first three of these factors were found to be most dominant in the decision making process. The final destination and proximity to the destination restricts the options for the ATCo to let the aircraft climb, descend or change heading. Each aircraft has an exit point and exit flight level, this is the point where the flight is handed over to the next sector. In many cases there are also procedures in place, agreements between the sectors, that further define the transfer to the next sector. These agreements for instance contain rules on the transfer flight level, or required minimum separation of flights. Lateral or vertical proximity to the transfer point limits the resolution options of the ATCo.

Another geographical factor that limits the resolution space is the occurrence of weather in the sector. A bad weather cell means that part of the airspace is not available, therefore the solution space that the ATCo can use becomes smaller.

The aircraft performance is a factor that can be used by the ATCo if he/she is familiar with the performance of an aircraft type or after requesting the climb/descent options from the pilot. The method that solves the conflict in the most efficient manner is also dependent on the time left to the predicted loss of separation. Solving a conflict with a small heading change is only possible when the change is initiated in a timely manner. Most conflicts are found the be solved at twelve to seven minutes before the loss of separation [18].

The controller's mental state can also impact the resolutions strategy. At a higher sector load there may not be sufficient time to find the most optimal solution for a conflict or, the ATCo may opt to take a resolution that requires less coordination effort.

3.5. Conclusion

Several insights from this description of the ATCo workdomain are relevant in the search for "basic" or "non-basic" traffic characteristics. This section summarises the factors that influence the perceived complexity.

First, for the resolution of any conflict the ATCo first tries to use a standard, frequently used, solution. The ATCo will only spend additional cognitive effort if a standard solution is not possible. This suggest that flights that are in a conflict for which a standard solution is available will be deemed less complex by the ATCo.

For conflict detection ATCos first look at the flight altitude. Flights that have sufficient vertical distance from other flight can never be in a conflict. As this information can simply be read from the flight label and it requires little effort to process, conflict detection for flights with sufficient vertical separation is easy.

Another factor that is shown to influence the effort it takes to detect and resolve a conflict is the angle of convergence between to flights. Both conflict detection and resolution are more complex for shallow convergence angles.

Finally, a flight with restricted manoeuvring space in a conflict has limited resolutions options. Which generally means that more cognitive effort is needed to find resolution for the conflict. The manoeuvring space can be limited due to the lateral and vertical distance to the sector boundary, the final destination, the proximity of the flight to a conflict, or a weather cell in the sector.

4
Complexity in air traffic

Complexity in air traffic has been defined in different ways. From the controller perspective, researchers attempted to find a measure for controller mental load to be able to understand the limitations of the human task demand load capacity, with the objective to increase the airspace capacity.

One approach is to define the task load in terms of sector complexity, based on sector and traffic characteristics. The difficulty of defining mental load in terms of the traffic characteristics is the influence of mediating factors. These factors form a bias between the complexity of the traffic scenario and mental load needed to manage the traffic.

Other models for air traffic complexity use the characteristics of the traffic or the sector to come up with a score for complexity. These models, or metrics, usually are meant to find a score for the complexity of all the traffic in a sector combined, instead of the complexity of a single flight. However, they are useful to the project to gain an initial inside into factors contributing the traffic complexity.

Section 4.1 discusses the controller mental load and mediating factors. Section 4.2 discusses different models for traffic or sector complexity. Section 4.3 proposes different factors from the existing models that can be useful for the determination of the complexity of an individual flight.

4.1. Controller mental load

In the search to learn the limits of controller capacity, different researchers have attempted to model the controller mental load [21]. Hilburn, in a review of studies into complexity for ATC, states that there is no widely accepted definition of mental load, however, most models are defined by a combination of factors relating to the difficulty of the task, the task demand load, and the controller's effort to complete the task, the workload or mental load.

The trouble in defining a controller mental load model comes from different shortcomings of modelling human behaviour. Most notably, it is difficult to objectively measure behaviour, as he cognitive processes inside of the brain cannot be observed. Also, measurements are often noisy and difficult to repeat due to individual differences between humans, and due to the fact that human behaviour is not constant but varies over time. The mental load, or effort perceived by the controller, has been found to be affected by his/her mental state at the time of the simulation. This makes difficult to reproduce results [21]. Human behaviour changes with time, because humans tend to use strategies, to lessen cognitive effort needed for performing a decision making task. ATCos have been shown to apply different strategies when confronted with scenarios with increasing traffic load [2]. The human controller does not always rationally minimise loss and maximise profit in a control system. Instead the decision making process is affected by making biased decisions when dealing with unpredictable events or high-risk situations. For example, human controllers tend to be less willing to take risks for profit than for losses. Moreover, human decisions are coloured by the availability of information from the working memory, and unwillingness to reconsider options for a given situations [56]. These strategies for decision making, that impact the perceived effort of task, develop with experience in performing a certain task. Thus, differences in perceived workload can be found between novice and experienced controllers.

To avoid the human controller bias, early mental load models only made use of the human's observable, physical, actions. These models would include for example the number of entries given by the controller or time spent in communication as a measure for workload, usually in combination with measures of complexity for the traffic scenario. Pawlak et al., however, argued that such is model is lacking, as the the physical work needed does not accurately reflect the mental load [35]. For example, vectoring an aircraft around a Special Use Airspace (SUA) includes multiple physical entries, but is usually a routine task for which low cognitive effort is required.

Hilburn et al. developed a model that relates task demand load to the mental load or workload. The task demand load is the objective level of difficulty of the task. For ATC task demand load is a combination of the the airspace and traffic characteristics [21], but also the concept of operation [29]. The concept of operation includes factors such as the required separation and the support tools available. In the model by Hilburn all factors that together are the task load are referred to as the system factors. The actual perceived workload is further affected by the operator factors, factors that are unique for each controller such as his/her skill level, experience and applied strategy.

A schematic overview of this model can be found in figure 4.1.

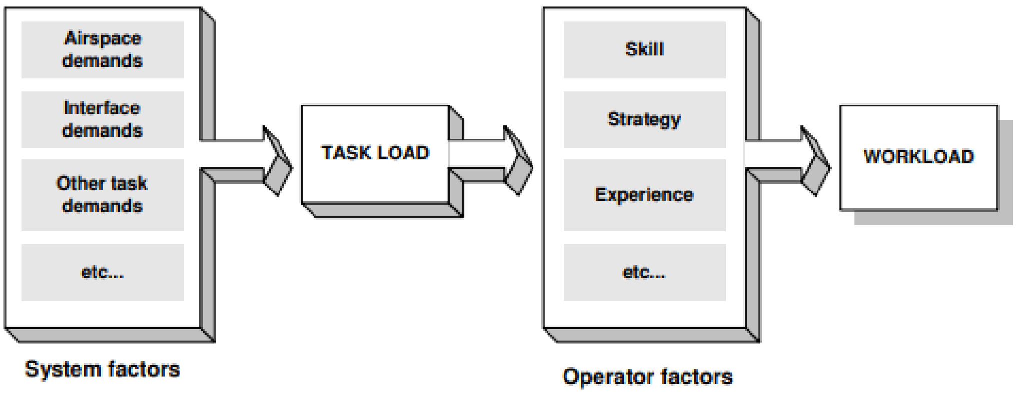


Figure 4.1: Relation between taskload and workload as presented by Hilburn et al. [21]

4.2. Complexity metrics

This section describes different metrics from literature that are use to determine the complexity traffic in a sector.

4.2.1. Aircraft Density (AD)

The AC metric is solely based on the number of aircraft present in a sector. A higher number of aircraft to control means a higher complexity. This metric is currently used in European airspace to determine whether a sector should be split or can be merged [38]. If the number of aircraft is predicted to increase above the maximum capacity of the sector the sector is split.

4.2.2. Dynamic Density DD

Many researchers argue that the complexity in the airspace, and consequently controller workload, depends on more factors than just the number of aircraft in the sector. As an extension to the AD method the DD metric includes factors that describe the structural and flow characteristics of the airspace [28] [21].

The structural characteristics define the fixed airspace in terms of factors such as: terrain, the number of airways and the number of airway crossing. These factors do not change with time. The flow characteristics are factors that depend on time and include factors like: the number of aircraft, the relative speeds of aircraft, the closing rates between aircraft or the presence of a weather cell.

All factors have been found by workload rating experiments and by interviewing ATCos. The relative weight of the factors is found through subjective ratings from ATCos for different traffic scenarios. Different DD models can be found in literature, due to the dependence on sector geometry factors and the subjective way in which the factors are found. This also implies that the DD metric is sector dependent [38]. Kopardekar et al. performed a regression analysis on complexity ratings to find a selection of DD factors from different metrics that most accurately describe the controller workload for any sector. The factors from this metric are listed and detailed here [25]:

- **Aircraft count:** Any aircraft that comes through the sector has to be accepted, monitored and transferred. Thus, each additional flight increases the controller workload even if the flight does not have any interactions with other flights. More flights in the sector limit the solution space for flights that are a conflict, which implies more effort is needed for conflict resolution.
- **Number of aircraft/occupied volume of airspace:** A better way to express the available solution space is to divide the number of flights over the volume of the airspace. Complexity of the sector increases when there is less space per flight. This factors shows how the complexity of a traffic scenario is affected by the sector geometry.
- **Proximity of conflicting aircraft with respect to their separation minima:** Flights that come close to or will violate the separation minima will require additional attention from the ATCo.
- **Number of climbing or descending aircraft:** Climbing or descending flights require extra monitoring as the change in flight level might lead to new conflicts in the traffic scenario. The level change is not visible on the 2D image of the traffic scenario adding an extra difficulty in the monitoring of these flights.
- **Time-to-go to conflict:** A shorter time to the conflict means less time for the ATCo to resolve the conflict adding to the complexity.
- **Ratio of standard deviation of speed to average speed:** More cognitive effort is needed to determine whether two flights on converging tracks will get into a conflict is the flights are travelling at different speeds [24].
- **Conflict resolution difficulty based on crossing angle:** The detection difficulty of a conflict is affected by the conflict angle between the crossing flights. For a 90 degrees crossing angle detection is easiest, as the angle moves away from 90 degree detection is harder for the ATCo [21].
- **Number of aircraft with a 3D Euclidean distance between 0-5 nautical miles excluding violations:** ATCos constantly monitor the traffic for potential conflicts, flight that come near the violation minima require more attention.
- **Number of aircraft with a 3D Euclidean distance between 5-10 nautical miles excluding violations:** ATCos constantly monitor the traffic for potential conflicts, flight that come near the violation minima require more attention.
- **Count of number of aircraft within a threshold distance to the sector boundary:** The proximity to the sector boundary limits the resolution options for a conflict, which increases the complexity.

4.2.3. Input-Output (IO)

The IO approach presented by Lee et al. defines complexity as the control activity needed to avoid interaction when a new aircraft enters a sector. Instead of looking at the characteristics of the traffic in the sector, the input-output approach looks at all the traffic in the sector as a system. The complexity is defined as the output from this system for a given input.

In the IO approach, the sector is a defined geographical area in the airspace. This is not necessarily an area with the geographical boundaries of a sector controlled by an ATCo. Instead it could be an area positioned around any number of flights. This definition makes the IO approach relevant for future TBO concepts where flights are no longer controlled in defined geographical areas. The system can also be extended with additional variables such as weather in the sector or the closure of part of the sector boundary.

The output of the system, the complexity, is the total heading change of all aircraft currently in the sector needed solve the system with the addition of one extra aircraft to the sector. To solve the system, each aircraft can get one heading change to ensure no separation minima will be violated. The aircraft travel at constant speed.

The aircraft entering the sector is the input of the system. The aircraft can enter the sector with any bearing and at any point of the sector boundary. The entry position is defined by the angle that the position makes with the sector centre. The entry bearing is the angle between the track of the aircraft and the line connecting the centre of the sector to the aircraft. Figure 4.2 illustrates the entry bearing and position definitions.

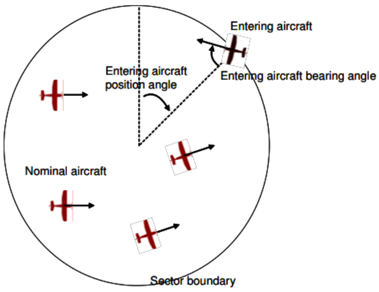


Figure 4.2: Aircraft position and bearing definitions [26]

The result of the IO approach is a complexity map that gives the complexity for an entering aircraft for a combination of position and heading. Figure 4.3 shows a sector scenario and the corresponding complexity map. For any combination of bearing and position this map gives a complexity score. Lee et al. suggest that the complexity map can be transformed to a measure for the total sector complexity using different methods each with a different outcome. Two methods suggested are either taking the worst complexity value in the map, or measuring the share of the map which each the complexity values is above a certain threshold [26].

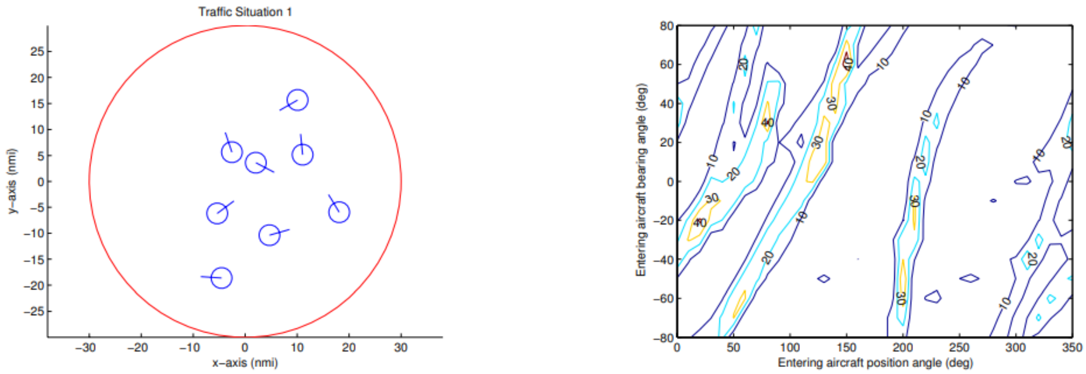
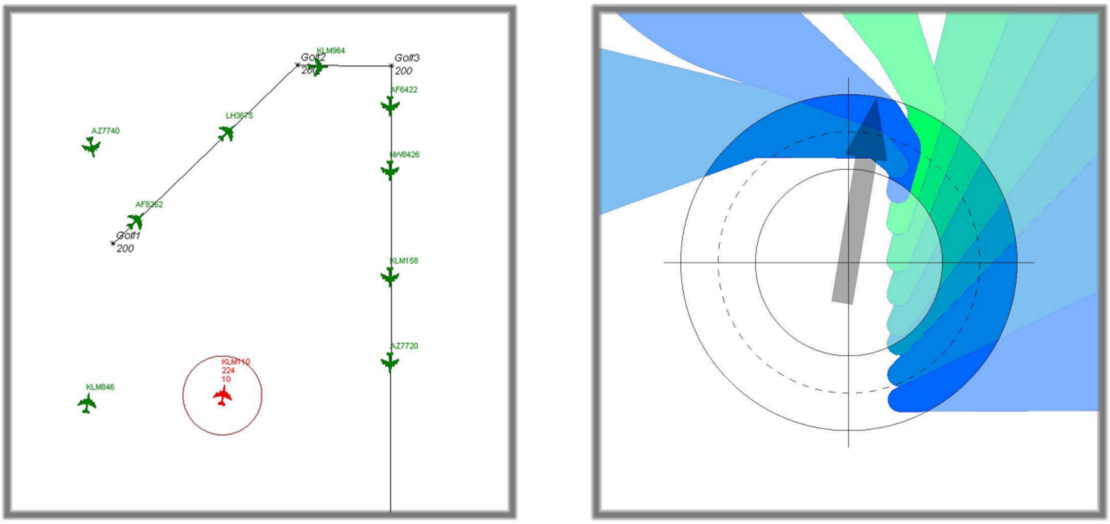


Figure 4.3: Complexity map from the input-output metric [26]

4.2.4. Solution Space

The Solution Space method relates the solution space around the track of an flight to the complexity of the sector. The solution space around an aircraft is defined as the combination of heading and speed clearances that can be issued by the controller that do not lead to a violation of separation minima. This methods thus relates the complexity of a traffic scenario to the position, direction and velocity of the aircraft in the sector [20]. The solution space is constructed from so called Conflict Zones around an aircraft. Each Conflict Zone represents the the headings and velocity combinations that lead to loss of separation with one other aircraft in the sector [57].

Figure 4.4 shows the construction of the solution space.

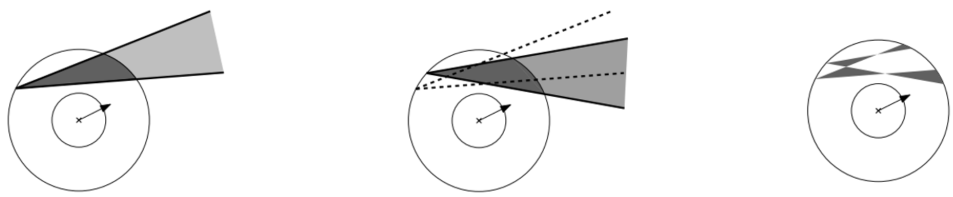


(a) An air traffic situation in Cartesian space

(b) A solution space Diagram with multiple CoZ's

Figure 4.4: The Solution Space diagram with Conflict Zones for other aircraft in the sector [57]

Two parameters are measured from the visual representation of the solution space. The first is the Area parameter, which is the percentage of the total area covered by Conflict Zones or the percentage of heading/velocity combinations that lead to loss of separation. The total of the solution space can be limited by changing the observation angle around the current track of the aircraft, to reduce the influence of solution space that is so far of the current track that it does not pose of feasible option for conflict resolution as it would be deemed too inefficient by the ATCo. Second, is the Change parameter, a time dependent parameter, that measures the percentage of solution space that changes from safe to unsafe or visa versa. For this parameter the solution space has to be recalculated over fixed time intervals [57]. Figure 4.5 illustrates the concept of the Change parameter.



(a) Initial solution space

(b) Next solution space calculation

(c) Resulting "Changed" area

Figure 4.5: Visualisation of the Change parameter [57]

The Solution Space metric has been found to be a robust method that performs well for workload prediction and is less sector dependent than methods like the DD metric [40]. However, this method does not consider effect of reduced or increased efficiency of the flight path to the sector exit. While the area of observation can be limited around the current track of the aircraft the Solution Space method does not make any distinction in the safe or unsafe areas. A flight with a high percentage of unsafe areas for which the most optimal path to the exit point of the flight is over a safe area would thus be considered complex.

Outcomes from the Solution Space parameters can be tuned by changing the extent of the used intent information of flights, or putting a limit on the time frame for which the Conflict Zones are calculated.

4.2.5. Intrinsic Methods

Due to the difficulty to mathematically model the subjective workload factors several metrics have been designed that disregard the effect of workload on complexity, instead these measures are based on the intrinsic traffic disorder. Delahaye et al. have described two different approaches based on this concept. The first metric uses the relative position and speed of each aircraft in the sector to determine the structure and disorder of the air traffic. Extending this work further, they define complexity by the topological entropy of the dynamic system that is a model of the air traffic [14].

4.3. Conclusion

Theory of controller workload shows that perceived workload is affect by many different factors, including procedures, equipment and individual cognitive differences. Even so different metrics have been developed that relate the complexity of the traffic scenario to the controller workload. In line with this approach this thesis project aims to define the complexity of a single aircraft in terms of the factors from the sector complexity metrics.

The AD, DD and intrinsic methods generally describe the complexity in terms of instantaneous traffic characteristics, such as the interaction geometries and factors that effect the effort of conflict detection.

The Solution Space metric and Input-Output approach are different because they additionally look at the effort of resolving interactions between flights.

For the complexity of a single flight both approaches can be relevant. Therefore, factors from the different metrics described are considered and compared in the development of a metric for a single flight.

5

Complexity metric design

Three different metrics for complexity of an individual flight have been formulated based on the information found in the literature study and after discussions with ATCos on air traffic complexity [1]. This chapter discusses the definitions, possible advantages or shortcomings of each of the metrics and the possible correlation between factors used. The metrics are (partly) adapted from the sector complexity metrics discussed in Chapter 4 to calculate a value for the complexity of a single flight.

5.1. Dynamic Density for a single flight

The DD metric for sector complexity comprises a number of factors that unevenly contribute to the total complexity. The relative importance of each of these factors is usually found from subjective ATCo ratings. To use the dynamic density principle for complexity of a single flight factors have been identified that can be used to describe the geometric properties of the interactions of a single flight with each of the other flights in the scenario. Initially, each of these factors is used individually, as a metric, to learn how they contribute to the perceived complexity. If from the experiment and discussion with the ATCos the relative importance of each of these factors becomes clear, a combined metric can be defined to find a single value for complexity. From an initial discussion with two MUAC ATCos[1] it is already noted that the minimum distance between two flights without vertical separation is the most important factor in the process to determine which flight require additional monitoring or action. The following factors are included in the DD metric:

Closest point of approach (CPA) horizontal:
The smallest distance between two flights during the course of the scenario in the horizontal plane.

Flights that are more than 1,000 ft apart vertically but have loss of separation in the lateral plane are not in conflict. However, a loss of separation in horizontal plane can create additional work, as the flights can clutter on the radar screen.

Closest point of approach 3D:
The smallest horizontal distance between two flights during the course of the scenario that are within 1,000 ft of each other vertically. For flights that are vertically separated no value is returned.

If the value is within the lateral violation minima, there is an interaction with loss of separation between the two flights. Such a violation will always require a corrective action from the ATCo. However, if the minimum distance is close to the violation limit, action is usually also taken to ensure separation. MUAC ATCos indicated that they take action in case the predicted minimum distance is below 7 NM. For a distance between 7 and 10 NM the ATCos allocate additional attention to the flight.

Crossing angle:
The crossing angle is the angle between two crossing flights. The angle is calculated for any two crossing flight paths that have no vertical separation.

The crossing angle affects both the ease of detection and resolution. Interactions with crossing angles away

[1] Discussion with two MUAC ATCos about interaction complexity on 14-02-2022

from 90 degrees are harder to detect, also the cognitive work to predict the minimum distance between the two flights becomes higher. Solving an interaction between two flights with shallow crossing angle is more complex as the solution space is limited, thus a vertical solution is often required. The vertical solution adds an extra dimension to the problem, adding to the complexity.

Time to CPA:
The time to CPA is defined as the time to go until a flight reaches point where the distance to another flight is minimal. The time to CPA is calculated for any two flight paths that have no vertical separation.

If the interaction is still far away there is added uncertainty in the prediction of possible loss of separation, which can mean additional attention from the ATCo is required. Conversely, a short time to a conflict increases the stress experienced by the ATCo as he/she has limited time to find a solution. In a discussion with MUAC ATCos they indicated that the look-ahead time in MUAC airspace largely depends on the sector. In some sectors ATCos look for interactions between flights that are 10-15 minutes away. In smaller sectors, such as the Brussels sector, this look-ahead time is around 5 minutes. Despite these large differences they did indicate that an loss of separation occurrence that is unsolved and less than 3 minutes away generally makes them nervous.

Duration of loss of separation:
This value is the duration for which two flights are within a defined distance of each other. It is calculated for any two crossing flight paths that have no vertical separation.

There is a correlation between this duration and the interaction geometry. A smaller minimum distance between two flights and a shallow crossing angle increase the loss of separation time. Another factor that affects this measurement is the (relative) speed of the aircraft.

Number of interactions:
To find the number of interactions that a flight encounters during the scenario an interaction is defined as flights coming within a specific distance. This specific distance can be varied to tune the outcome of the measurement.

The number of interactions is affected solely by the CPA between two flights. The other geometry factors, such as the crossing angle, are not reflected in this measurement. This measure thus might not fully reflect the complexity of different interactions on the route of the flight. However, from the simulation data it might be possible to find a relation between complexity and the number of flights interacting. Varying the threshold distance can give an insight in the specific interaction distance that is considered complex by the ATCos. Important distances to consider are those indicated as complex by the ATCo, as mentioned above, all interactions with minimum separation under 10 NM. However, another interesting find could be a threshold for the minimum distance above which the flight is generally not given additional attention by the ATCos.

5.2. Solution Space metric

The Solution Space metric is a measure for the solution space available for a flight, this metric is the Solution Space Area parameter as defined in Section 4.2.4. The geometry of the traffic scenario and the interactions affect the solution space. The Solution Space Area parameter thus is a metric that combines the metrics mentioned above related to the interaction geometry. While indicating the possible space for resolution around the flight path, the need for a resolution is not reflected in the Solution Space metric. Flights that are on an interaction free flight path may still receive a high complexity rating from this metric.

While this metric has been shown to be useful for the complexity measurement of the complete sector it is unclear if the metric is equally useful for the assessment of complexity of a single flight. In the Solution Space Area the percentage of blocked solution space is affected by the geometric properties of the interactions. Comparing this metric to DD metric might give an insight into geometric properties of an interaction that have an additional or reduced contribution to the perceived complexity. For instance, a more shallow crossing angle gives a larger blocked area of the Solution Space Diagram, however, from the crossing metric a threshold value may be found below which the effect on perceived complexity is significantly larger, such a steep increase is not reflected in the Solution Space metric. Resolutions proposed by ARGOS are subject to larger separation minima, to create more stable solutions. ATCos probably also consider this a factor when finding a resolution. The separation minima used to calculate the blocked areas of the Solution Space diagram can be varied to find a value for which it gives the best indication for complexity.

5.3. Resolution metric

The resolution metric calculates the complexity of simple solutions. These solutions can be a resolution for an interaction on the current flight path, or an optimisation of the current flight path. The solutions considered are based on the objectives for solutions defined by the ARGOS logic, as described in Chapter **??**. Table 5.1 describes the different solutions used in the Resolution metric.

Table 5.1: Overview of the solution options for the Resolution metric

Solution	**Description**
Current route @ AFL	The current flight path as described in the flight plan keeping the current flight level.
Current route @ TFL	The current flight path at the transfer flight level.
Current route +- small vertical change from AFL	The current flight path with a small flight level change from the current flight level.
DCT @ AFL	Direct to the sector exit point keeping the current flight level.
DCT @ TFL	Direct to the sector exit point at the transfer flight level.
DCT +- small vertical change from AFL	Direct to the sector exit point with a small level change from the current flight level.

The complexity of solutions is defined as the total number of interaction on a track. Again the number of interactions found can be regulated by setting a threshold for a minimum distance between two flights, under which the interaction counts towards the metric. The objective of the resolution metric is to determine the effect that the availability of simple resolutions has on the perceived complexity. It does, in a way, look at the solution space of a flight, however, different than for the Solution Space metric, only commonly used or optimised solutions are considered. The measure for complexity of a track could also be extended to include other geometric properties of the interactions. If the relative importance of different geometric properties is determined, it might be possible to make a more accurate prediction of the complexity of a resolution.

5.4. Conclusion

The three proposed metrics describe the complexity in terms of the interaction geometry, solutions space and availability of relevant resolutions. To find a correlation between perceived complexity and the metric outcomes an experiment has to be performed to gather data on the perceived complexity of an ATCo for an individual flight in a traffic scenario. Using these data the metrics can be further developed by adapting the look-ahead time or measurement range.

6
Preliminary experiment design

The objective of the designed experiment is to identify factors, related to the traffic characteristics, that can be used to describe the complexity of a single flight in a traffic scenario. Ideally, from the experiment data a combination of factors are found that can be combined in a final metric for complexity. To increase the usability of this metric the input of the experiment are factors related to the geometry of the traffic in the scenario disregarding the contribution of the sector geometry on complexity. The following research question is used as a basis for the experiment:

What traffic characteristics are the most relevant to determine the perceived complexity of an additional flight entering a traffic scenario?

In the experiment MUAC ATCos are asked to rate the complexity of a single flight entering different traffic scenarios. Different scenarios are designed to get a complexity rating from the ATCos for flights with different characteristics. The rating and logged flight data will be used to find a correlation between the outputs of different metrics, as described in Chapter 4.2, and the perceived complexity.

6.1. Participants and apparatus

MUAC ATCos will participate in the experiment. Their standard radar screen is mimicked with a TU Delft-built Java-based simulator, reducing the time needed for the ATCos to familiarise themselves with the interface. Different from the real ATCo workdomain clearance issues can only be given via CPDLC, as there is no radio connection. The simulator is controlled using a computer mouse and keyboard.

6.2. Procedure and tasks

The experiment starts with briefing and training phase. The briefing includes a introduction of work done on single flight complexity. A preview of the lay-out of the experiment scenarios can also be included. During the training phase the ATCos are introduced to the simulator environment and can familiarise themselves with the differences between their real workdomain and the simulator in managing the aircraft. During the experiment multiple traffic scenarios are presented to the ATCos. In each scenario one flight is highlighted as the flight of interest, this flight is at the sector boundary, entering the sector. For each of the scenarios they are given time to study the traffic. After they have familiarised themselves with the scenario they are asked a for subjective rating of the complexity of the flight of interest. Next, they are asked to resolve the interaction by issuing instructions only to the highlighted flight. This step is included to assess the usability of the easy resolutions to determine flight complexity. Scenarios may be presented for which the ATCo determines there is no interaction that needs to be resolved. This can occur for instance of the predicted CPA is close to the loss of violation minima but the ATCo decides to take no action. After all scenarios have been viewed the ATCos are asked to provide additional comments on the flights that they rated least and most complex. This discussion should provide additional insight in what factors they considered in their rating of the flight.

6.3. Independent variables

The independent variables are the design variables of the traffic scenarios. The designs can be altered with the number of aircraft present and the positions, direction and speed of the aircraft. The scenarios should be designed such that a relation can be found between the metric outputs and the complexity. For the scenarios real traffic patterns provided by MUAC are used. The aircraft of interest is added to this scenario at a location close to the sector boundary, when the aircraft is entering the sector. ATCos have different work styles depending on the sector that they are licensed for[1], which might affect which flights they perceive as complex. Ideally, to find a more generalised definition of single flight complexity, sectors from each of the sector groups of MUAC are used.

In the design the effects of correlation between the metrics should be accounted for. For the metrics related to the interaction geometry interactions, should be designed that are varied for only one geometry factor, while the other factors are kept constant. For example for an interaction with the same crossing angle and time to the CPA point, the CPA distance can be varied to find the effect of the CPA on complexity. The strategy is applicable for changes to the CPA, crossing angle and time to CPA by altering the aircraft heading, speed and initial position in the scenario. There is a correlation between the duration of the loss of separation and the CPA, crossing angle and speed of the aircraft involved in the interaction. The correlation between the crossing angle and the duration of the loss of separation is illustrated in Figure 6.1. For this metric, interactions can be designed that have the same loss of separation time but a different interaction geometry, to find the effect of the duration of the loss of separation on the perceived complexity.

In the Solution Space metric, the interaction geometry and space for resolutions have a combined effect. To assess the applicability of this metric for single flight complexity, scenarios should be designed in which the aircraft of interest has a similar solution space availability as a result of varying interaction locations and geometries. In this manner the relative importance of the interaction geometry and availability of solutions can be found.

To find the influence of the availability of easy solutions to the interactions similar scenarios can be compared for which the availability of the easy solutions is varied. This could be done by blocking solutions using other traffic, or a weather or forbidden zone. Other possibility is simply imposing restrictions on the options available to the ATCo for resolutions. Using this same approach the effect of a possibility to optimise the flight path on the complexity score can be assessed.

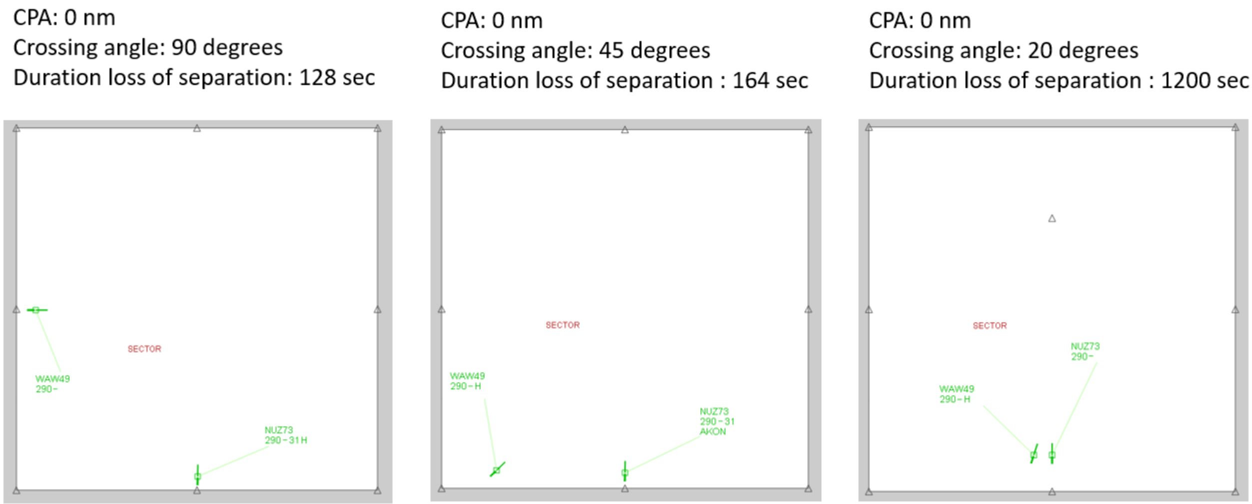


Figure 6.1: Visual representation correlation between crossing angle and duration loss of separation

[1]Discussion with two MUAC ATCos about interaction complexity on 14-02-2022

6.4. Dependent measures

The dependent measures are the data collected from the experiment and the output of the metric used to assess the complexity. The complexity metrics log a complexity value for the traffic situation at each radar update.

The dependent measures from ATCos data are:

1. **Complexity rating:** For each scenario the controllers are asked to rate the complexity of the flight of interest on a slider scale.
2. **Resolution data:** The issued resolution clearances.
3. **Additional questionnaire:** Each participants is asked to fill out an additional questionnaire about the flights that have been rates most and least complex. In this questionnaire the participants are asked to comment on the factors that were most relevant in determining the complexity score.

The metric output variables are:

1. **CPA 2D:** The CPA for the flight of interest for all flights in the scenario determined in the lateral plane only.
2. **CPA 3D:** The CPA for the flight of interest with for all flights in the scenario for which there is a vertical violation.
3. **Crossing angle:** The crossing angle between the flight path of the flight of interest and all other flights in the scenario for which there is no vertical separation.
4. **Time to CPA:** The time remaining until the flight of interest reaches its CPA with any of the other flights in the scenario for which there is no vertical separation.
5. **Duration of loss of separation:** The total time for which the flight of interest and any other flight in the scenario are with a specified distance of each other, when there is no vertical separation.
6. **Number of interactions:** The number of interactions where the CPA is below a specific distance and there is no vertical separation on the cleared flight path for the flight of interest.
7. **Solution Space Area parameter:** The percentage of blocked space on the SSD for the flight of interest.
8. **Resolution metric data:** The number of interactions on the relevant easy solutions for the flight of interest.

7 Conclusion

Increasing levels of automation will, in the coming decades, change the working of the ATM sector, and by extension the ATCo work. For the introduction of automation assistance different concepts have been developed. This thesis project concentrates on the concept of mixed responsibility between ATCo and automation in the handling of traffic. Recent research has reviewed the effects and shown the feasibility of this shared responsibility concept [7], [13]. To successfully implement this concept it has been proposed to allocate the "basic" traffic to the automation system [46], [13].

A first step in the allocation of "basic" flights is to identify flight characteristics that contribute to perceived flight complexity. To that end a literature study was performed to find the currently used methods that define air traffic complexity. The literature study found that the task complexity coming from the air traffic and sector characteristics is not the only factor factor contributing to ATCo workload. Individual differences between controllers, ATC procedures and the tools available and other operator factors also affect the perceived workload. However, efforts to find a generic method to assess controller workload resulted in metrics based on traffic and sector geometry. The objective of these metrics has been to rate the total workload of managing all flights in a traffic scenario bound by the geographical sector boundaries.

From the existing complexity metrics and the study of the ATCo workdomain, flight characteristics have been found that are expected to affect the complexity of an individual flight. These characteristics have been combined in three metrics, the metrics describe the complexity for an individual flight based on either the interaction geometry or the solution space available.

In the next stage of this thesis project a simulation experiment will be performed. During this experiment ATCos will be asked to rate the complexity of a single flight in different traffic scenarios. The objective of this experiment is to find a correlation between to output of the complexity metrics and the subjective controller data. The results of this experiment should provide an insight into what flight (path) characteristics best reflect the complexity of a single flight.

Bibliography

[1] Discussion with two MUAC ATCos about interaction complexity, 2 2022.

[2] Sylvie Athènes, Philippe Averty, Stephane Puechmorel, Daniel Delahaye, and Christian Collet. ATC complexity and controller workload: Trying to bridge the gap. *Proceedings of the International Conference on HCI in Aeronautics*, (56), 2002.

[3] Aviation24.be. Become an air traffic controller at Eurocontrol, Maastricht, 2017. URL `https://www.aviation24.be/air-traffic-control/eurocontrol/become-an-air-traffic-controller-at-eurocontrol-maastricht`.

[4] J. Elin Bahner, Anke Dorothea Hüper, and Dietrich Manzey. Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human Computer Studies*, 66(9), 2008. ISSN 10715819. doi: 10.1016/j.ijhcs.2008.06.001.

[5] L. Bainbridge. IRONIES OF AUTOMATION. pages 129–135, 1983. ISBN 0080293484. doi: 10.1016/b978-0-08-029348-6.50026-9.

[6] Bruno Berberian, Oliver Ohneiser, Francesca de Crescenzio, Fabio Babiloni, Gianluca di Flumeri, and Andreas Hasselberg. MINIMA project: Detecting and mitigating the negative impact of automation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10275 LNAI, 2017. doi: 10.1007/978-3-319-58472-0{_}8.

[7] Bettina Birkmeier, Sebastian Tittel, and Bernd Korn. Controller team possibilities for sectorless air traffic management. In *ICNS 2016: Securing an Integrated CNS System to Meet Future Challenges*, 2016. doi: 10.1109/ICNSURV.2016.7486362.

[8] A. Bisseret. Analysis of mental processes involved in air traffic control. *Ergonomics*, 14(5), 1971. ISSN 13665847. doi: 10.1080/00140137108931276.

[9] Bill Booth. Consolidated Validation Prototyping Report on Queue, Trajectory and Separation Management . Technical report, Eurcontrol, 2009. URL `https://www.eurocontrol.int/sites/default/files/library/014_Queue_trajectory_and_separation_management.pdf`.

[10] Carmen Bruder, Hinnerk Eißfeldt, Peter Maschke, and Catrin Hasse. A Model for Future Aviation. *Aviation Psychology and Applied Human Factors*, 4(1), 2014. ISSN 2192-0923. doi: 10.1027/2192-0923/a000051.

[11] Maia B. Cook, Cory A. Rieth, and Mary K. Ngo. Displays for effective human-agent teaming: The role of information availability and attention management. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9179, 2015. doi: 10.1007/978-3-319-21067-4{_}19.

[12] Nelson Cowan. Activation, attention, and short-term memory. *Memory & Cognition*, 21(2), 1993. ISSN 0090502X. doi: 10.3758/BF03202728.

[13] Gijs de Rooij, Clark Borst, Marinus M Van Paassen, and Max Mulder. Flight Allocation in Shared Human-Automation En-Route Air Traffic Control. 1 2021.

[14] Daniel Delahaye and Stephane Puechmorel. Air traffic complexity: towards intrinsic metrics. *3rd USA/Europe Air Traffic Management R&D Seminar*, (June), 2000.

[15] M. R. Endsley and E. O. Kiris. The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 1995. ISSN 00187208. doi: 10.1518/001872095779064555.

[16] Mica R. Endsley. Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting*, 32(2), 1988. ISSN 0163-5182. doi: 10.1177/154193128803200221.

[17] Mica R. Endsley. From Here to Autonomy: Lessons Learned from Human-Automation Research. *Human Factors*, 59(1), 2017. ISSN 15478181. doi: 10.1177/0018720816681350.

[18] Klaus Eyferth, Cornelia Niessen, and Oliver Spaeth. A model of air traffic controllers' conflict detection and conflict resolution. *Aerospace Science and Technology*, 7(6), 2003. ISSN 12709638. doi: 10.1016/S1270-9638(03)00064-6.

[19] Marc Christopher Gelhausen, Peter Berster, and Dieter Wilken. Post-COVID-19 Scenarios of Global Airline Traffic until 2040 That Reflect Airport Capacity Constraints and Mitigation Strategies. *Aerospace*, 8 (10), 2021. ISSN 22264310. doi: 10.3390/aerospace8100300.

[20] P. Hermes, M. Mulder, M. M. Van Paassen, J. H.L. Boering, and H. Huisman. Solution-space-based analysis of the difficulty of aircraft merging tasks. *Journal of Aircraft*, 46(6), 2009. ISSN 15333868. doi: 10.2514/1.42886.

[21] B Hilburn. Cognitive Complexity in Air Traffic Control a Literature Review. *Eurocontrol*, 12(13), 2004. ISSN 0014-0139.

[22] Christian P. Janssen, Stella F. Donker, Duncan P. Brumby, and Andrew L. Kun. History and future of human-automation interaction. *International Journal of Human Computer Studies*, 131, 2019. ISSN 10959300. doi: 10.1016/j.ijhcs.2019.05.006.

[23] K. Wolfgang Kallus, Dominique Van Damme, and Andrea Dittmann. Integrated Task and Job Analysis of Air Traffic Controllers - Phase 2 : Task Analysis of En-route Controllers. *Eurocontrol*, 1999.

[24] K. A. Kimball, M. A. Hofmann, and R. O. Nossaman. Differential velocity and time prediction of motion. *Perceptual and motor skills*, 36(3), 1973. ISSN 00315125. doi: 10.2466/pms.1973.36.3.935.

[25] Parimal H Kopardekar and Sherri Magyarits. Measurement and prediction of dynamic density. *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar*, (Dd), 2003.

[26] Keumjin Lee, Eric Feron, and Amy Pritchett. Air traffic complexity: An input-output approach. In *Proceedings of the 7th USA/Europe Air Traffic Management Research and Development Seminar, ATM 2007*, 2007.

[27] Mikell P. Groover. Automation, 2018. URL `https://www.britannica.com/technology/automation`.

[28] Mogford, R. H., Guttman, J. A., S. L. Morrow, and P. Kopardekar. The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature. Technical Report 3, 1995.

[29] NAT. Complexity Metrics for ANSP Benchmarking Analysis. Technical Report April, 2006.

[30] C. Niessen and K. Eyferth. A model of the air traffic controller's picture. *Safety Science*, 37(2-3):187–202, 3 2001. ISSN 09257535. doi: 10.1016/S0925-7535(00)00048-5.

[31] R. Parasuraman. Human-computer monitoring. *Human Factors*, 29(6), 1987. ISSN 00187208. doi: 10.1177/001872088702900609.

[32] R Parasuraman, T Bahri, J. E Deaton, J. G Morrison, and M Barnes. *Theory and design of adaptive automation in aviation systems.* 1992.

[33] Raja Parasuraman, Robert Molloy, and Indramani L. Singh. Performance Consequences of Automation-Induced "Complacency". *The International Journal of Aviation Psychology*, 3(1), 1993. ISSN 15327108. doi: 10.1207/s15327108ijap0301{_}1.

[34] Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, 30(3), 2000. ISSN 10834427. doi: 10.1109/3468.844354.

[35] William S. Pawlak, Christopher R. Brinton, Kimberly Crouch, and Kenneth M. Lancaster. A framework for the evaluation of air traffic control complexity. In *1996 Guidance, Navigation, and Control Conference and Exhibit*, 1996. doi: 10.2514/6.1996-3856.

[36] Linda Pfeiffer, Nicholas Hugo Muller, and Paul Rosenthal. A survey of visual and interactive methods for air traffic control data. In *Proceedings of the International Conference on Information Visualisation*, volume 2015-September, 2015. doi: 10.1109/iV.2015.102.

[37] Linda Pfeiffer, Georg Valtin, Nicholas Hugo Müller, and Paul Rosenthal. Aircraft in Your Head: How Air Traffic Controllers Mentally Organize Air Traffic. *HUSO 2015 : The First International Conference on Human and Social Analytics*, (c), 2015.

[38] Maria Prandini, Luigi Piroddi, Stephane Puechmorel, and Silvie Luisa Brázdilová. Toward air traffic complexity assessment in new generation air traffic management systems. *IEEE Transactions on Intelligent Transportation Systems*, 12(3), 2011. ISSN 15249050. doi: 10.1109/TITS.2011.2113175.

[39] Thomas Prevot, Jeffrey R. Homola, Lynne H. Martin, Joey S. Mercer, and Christopher D. Cabrall. Toward Automated Air Traffic Control-Investigating a Fundamental Paradigm Shift in Human/Systems Interaction. *International Journal of Human-Computer Interaction*, 28(2), 2012. ISSN 10447318. doi: 10.1080/10447318.2012.634756.

[40] Siti Mariam Abdul Rahman, Clark Borst, Max Mulder, and Rene van Paassen. Sector complexity measures: A comparison. *Jurnal Teknologi*, 76(11), 2015. ISSN 01279696. doi: 10.11113/jt.v76.5923.

[41] Esa M. Rantanen and Ashley Nunes. Hierarchical conflict detection in air traffic control. *International Journal of Aviation Psychology*, 15(4), 2005. ISSN 10508414. doi: 10.1207/s15327108ijap1504{_}3.

[42] R. W. Remington, J. C. Johnston, E. Ruthruff, M. Gold, and M. Romera. Visual search in complex displays: Factors affecting conflict detection by air traffic controllers. *Human Factors*, 42(3), 2000. ISSN 00187208. doi: 10.1518/001872000779698105.

[43] Ericka Rovira and Raja Parasuraman. Controller performance and attention allocation in future air traffic management: Effects of pilot intent information. In *Proceedings of the Human Factors and Ergonomics Society*, 2006. doi: 10.1177/154193120605000101.

[44] Francisco Javier Sáez Nieto. The long journey toward a higher level of automation in ATM as safety critical, sociotechnical and multi-Agent system. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 230(9), 2016. ISSN 20413025. doi: 10.1177/0954410015596763.

[45] Juergen Sauer, Chung Shan Kao, and David Wastell. A comparison of adaptive and adaptable automation under different levels of environmental stress. *Ergonomics*, 55(8), 2012. ISSN 00140139. doi: 10.1080/00140139.2012.676673.

[46] SESAR Joint Undertaking. ATM Master Plan. Technical report, 2020. URL `https://www.sesarju.eu/masterplan2020`.

[47] Steven T. Shorrock and Anne Isaac. Mental imagery in air traffic control. *International Journal of Aviation Psychology*, 20(4), 2010. ISSN 10508414. doi: 10.1080/10508414.2010.487008.

[48] SKYbrary. Area Proximity Warning.

[49] SKYbrary. Tactical Controller Tool (TCT), 2021.

[50] Norman J. Slamecka and Peter Graf. The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, 4(6), 1978. ISSN 0096-1515. doi: 10.1037/0278-7393.4.6.592.

[51] Oliver Späth and Klaus Eyferth. Conflict Resolution in En Route Traffic - A Draft Concept for an Assistance System Compatible with Solutions of Air Traffic Controllers. 2001.

[52] J. C. Sperandio. Variation of operators strategies and regulatingeffects on workload. *Ergonomics*, 14(5), 1971. ISSN 13665847. doi: 10.1080/00140137108931277.

[53] Mark St. John, David A. Kobus, Jeffrey G. Morrison, and Dylan Schmorrow. Overview of the DARPA Augmented Cognition technical integration experiment, 2004. ISSN 10447318.

[54] Barry Strauch. Ironies of Automation: Still Unresolved after All These Years. *IEEE Transactions on Human-Machine Systems*, 48(5), 2018. ISSN 21682291. doi: 10.1109/THMS.2017.2732506.

[55] Monica Tavanti and Matthew Cooper. Looking for the 3D picture: The spatio-temporal realm of student controllers. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5619 LNCS, 2009. doi: 10.1007/978-3-642-02806-9{_}122.

[56] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185 (4157), 1974. ISSN 00368075. doi: 10.1126/science.185.4157.1124.

[57] M. M. Van Paassen, Jurriaan G. d'Engelbronner, and Max Mulder. Towards an air traffic control complexity metric based on workspace constraints. In *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2010. doi: 10.1109/ICSMC.2010.5641823.

[58] Kim Phuong L. Vu, Thomas Z. Strybel, Vernol Battiste, Joel Lachter, Arik Quang V. Dao, Summer Brandt, Sarah Ligda, and Walter Johnson. Pilot Performance in Trajectory-Based Operations Under Concepts of Operation That Vary Separation Responsibility Across Pilots, Air Traffic Controllers, and Automation. *International Journal of Human-Computer Interaction*, 28(2), 2012. ISSN 10447318. doi: 10.1080/10447318.2012.634761.

[59] Earl L. Wiener and Renwick E. Curry. Flight-deck automation: Promises and problems. *Ergonomics*, 23 (10), 1980. ISSN 13665847. doi: 10.1080/00140138008924809.

A

Briefing Experiment

Introduction single flight complexity

Past research in air traffic complexity has mainly focused on the measurement of the overall sector complexity (e.g., counting the number of flights passing through the sector per time unit) and how that relates to self-perceived workload. This research will focus on the perceived complexity of a single flight and what objective factors contribute most to this self-assessed complexity. The possible outcomes of this study might have an impact on how and when flights should be allocated to automation or controllers (e.g., in a Flight Centric ATC environment).

Traffic scenario design

During the experiment, 36 traffic scenarios will be presented that include (slightly modified) traffic samples taken from a replay of traffic that travelled through the MUAC airspace. To this traffic a single flight is added, that will be the flight of interest for this experiment. The flight symbol and flight label of this flight are coloured blue. The “normal” traffic is conflict free, while the flights of interest may have a future conflict if no action is taken.

The position and heading of the flight of interest may in some cases be unrealistic, however, for the purpose of this experiment we ask you to assess the traffic situation as it is and attempt to ignore the fact that you are presented with a situation that does not usually occur. During the experiment you will be asked to assess the complexity for this flight of interest and indicate what other flights in the scenario played a role in making this assessment. The complexity in this case is defined as the attention that you think is needed to manage this flight. Thus, flights for which more attention and time will be required from you to find and execute a safe and efficient plan to bring this flight to the sector exit are considered more complex. Also, if the flight affects the rest of the sector plan it may be considered more complex.

Tasks and procedure

This section gives a short introduction of the tasks you will perform during the experiment. No further preparation is required.

1. The experiment starts with four training scenarios. In this training phase you can familiarise yourself with the controls of the simulator and the procedure to give the complexity score. The traffic presentation will be the same as in the rest of the experiment.
   Controlling the scenario: you can use your computer mouse to interact with the scenarios. By clicking and holding the next cleared heading in the flight label you can see the planned flight path. When you click on the aircraft ID in the flight label you can access a menu to use the VERA tool.
2. After the training phase 36 scenarios will be presented. For each of the scenarios you are asked to perform the following tasks:
   - The scenario is presented statically. You are asked to indicate which flights you considered when determining a complexity score for the flight of interest. You can indicate a flight by clicking on the flight ID in the flight label and clicking INCLUDE on the menu. If you want to exclude a flight you can click the menu again and select EXCLUDE. A blue frame will appear around the flight label of the indicated flight. You can find two figures of the experiment interface below.
   - After you have included the flights, you are asked to rate the complexity of the flight of interest by clicking on a the scale.
   - After giving a complexity score, the next scenario is automatically presented with a new flight of interest.
3. When you have completed all scenarios, you are asked to fill out an additional survey about the scenarios in which you gave the highest and lowest complexity score.

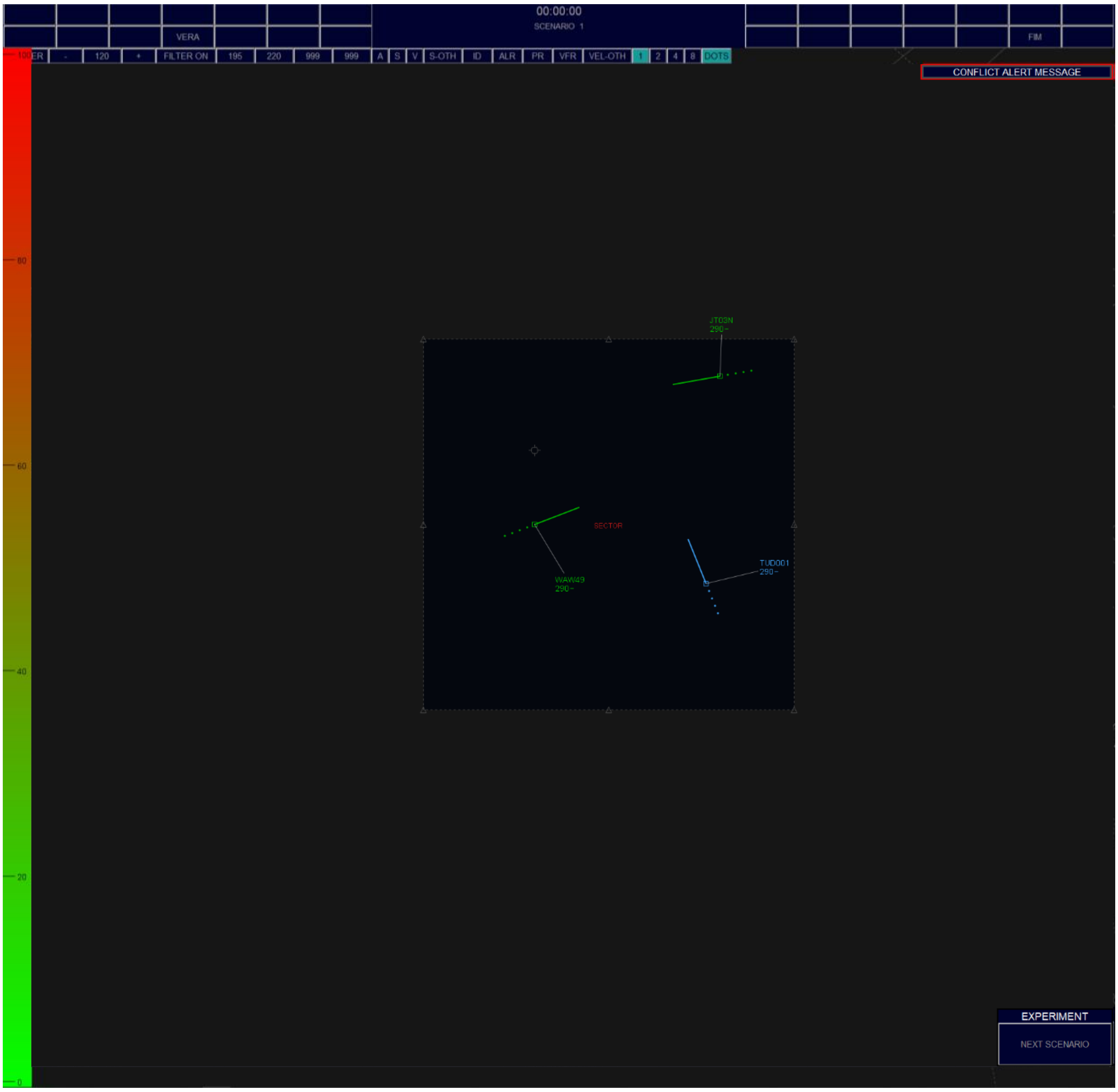

00:00:00
SCENARIO 1
VERA
FIM
FILTER ON
CONFLICT ALERT MESSAGE
JT03N
TUD001
WAW49
SECTOR
EXPERIMENT
NEXT SCENARIO

Figure 1

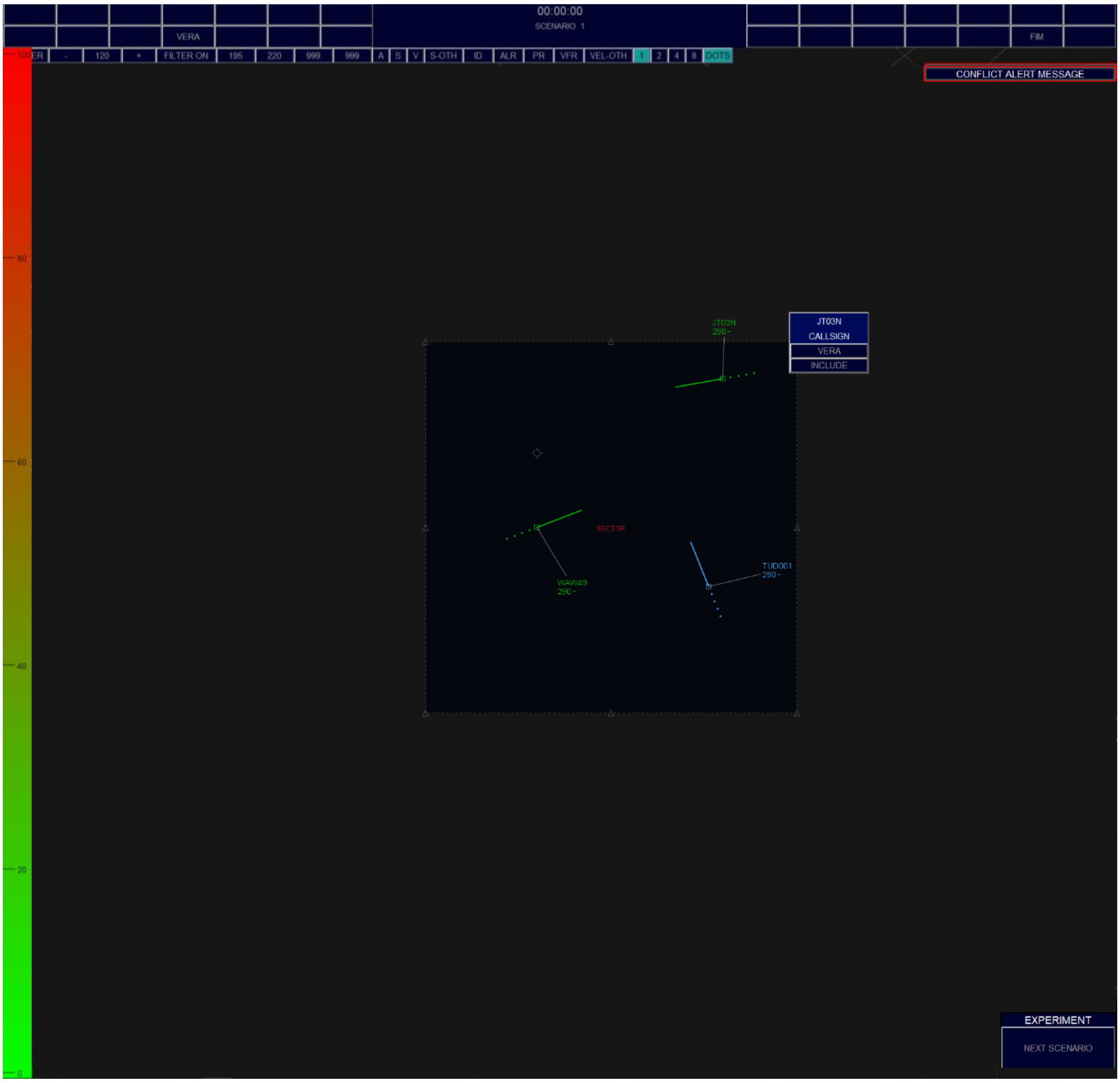

00:00:00
SCENARIO 1
VERA
FIM
FILTER ON
CONFLICT ALERT MESSAGE
JT03N
CALLSIGN
VERA
INCLUDE
SECTOR
WAW49
TUD001
EXPERIMENT
NEXT SCENARIO

Figure 2

B

Experiment Procedure

The experiment had three phases: training, experiment scenarios and the questionnaire. The training session started with one scenario with only two flights used to show the participants how to interact with the scenarios. Figure fig. B.1 shows the first training scenario for the Brussel scenarios. The training phase also had four scenario with the same background traffic used in the remainder of the experiment, which allowed the participants to familiarise themselves with this traffic. The background traffic was unique for each sector. Figures fig. B.2, fig. B.3 and fig. B.4 show the traffic scenarios for each sector.

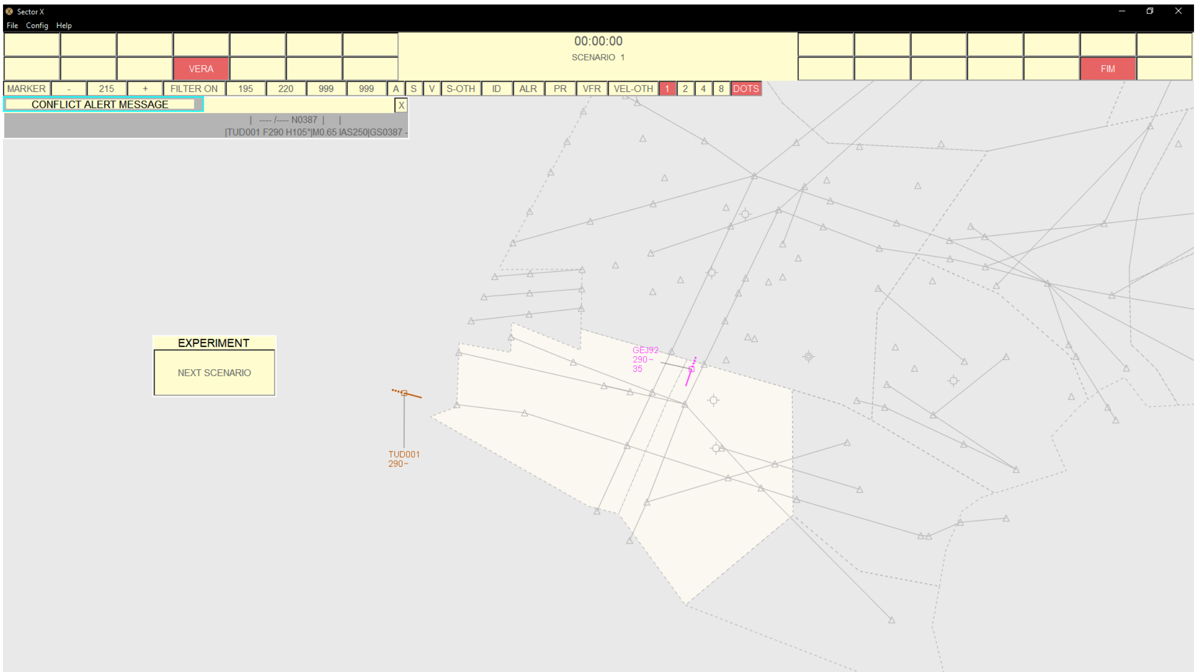


Figure B.1: Screenshot of the practise scenario for the Brussel sector

During the experiment the participants had two main tasks. First, they were asked to indicate flights from the background traffic that affected the complexity of the flight of interest. After doing this they could rate the complexity of the flight of interest, by clicking on the next scenario button. If no flights were included a message would appear asking if they were sure they did not want to included any flights (Figure fig. B.5). Figure fig. B.6 shows the 0-100 complexity rating bar on the left side of the radar screen. A small line indicated the score they give for the previous scenario.

In the final phase of the experiment a questionnaire was used to gain more insight in the factors that affect the complexity of nine scenarios. The questionnaire can be found in Appendix appendix C.

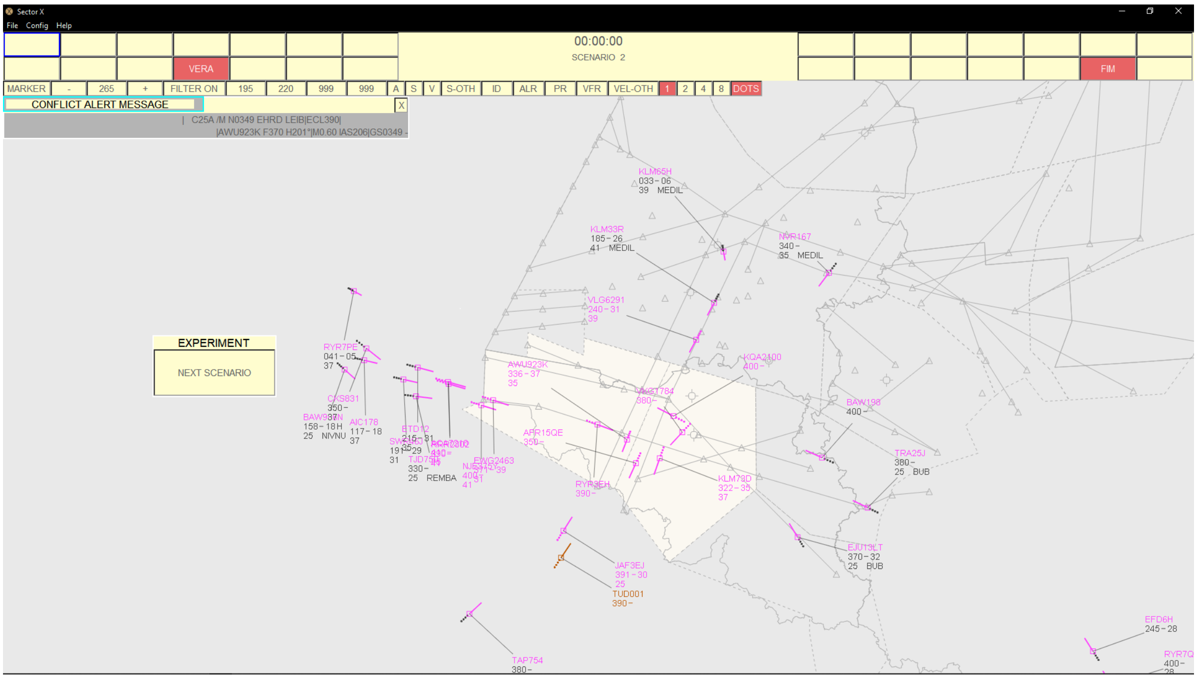


Figure B.2: Screenshot of a Brussel scenario

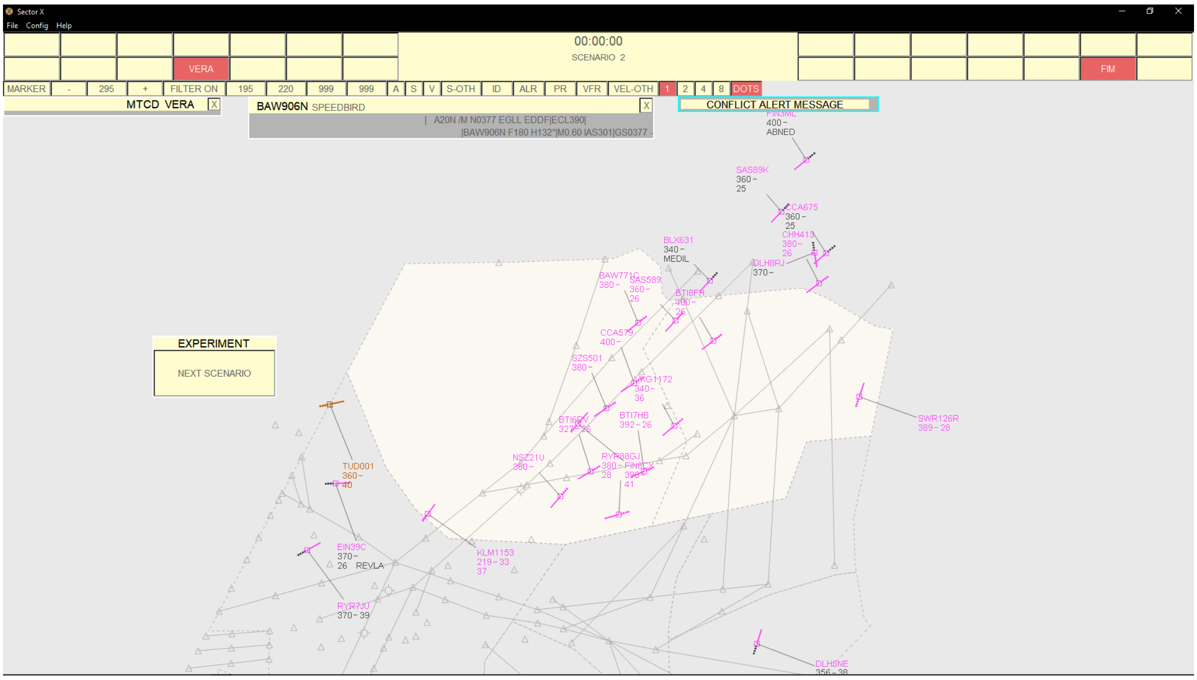


Figure B.3: Screenshot of a Jever scenario

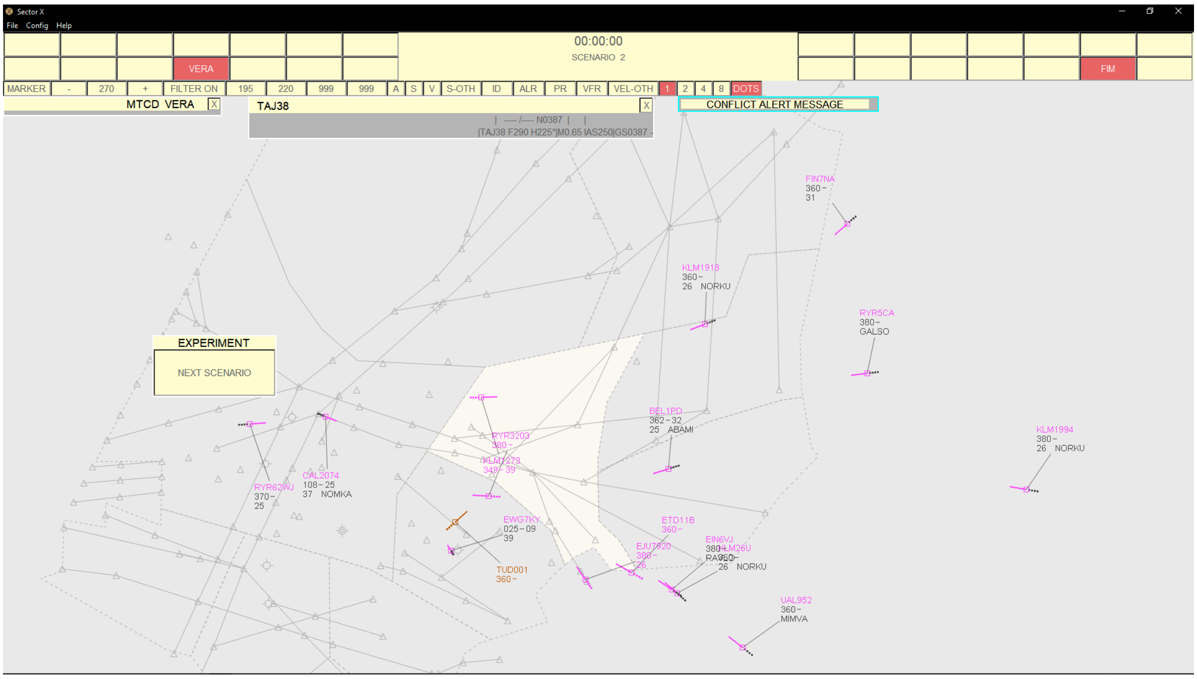


Figure B.4: Screenshot of a Munster scenario

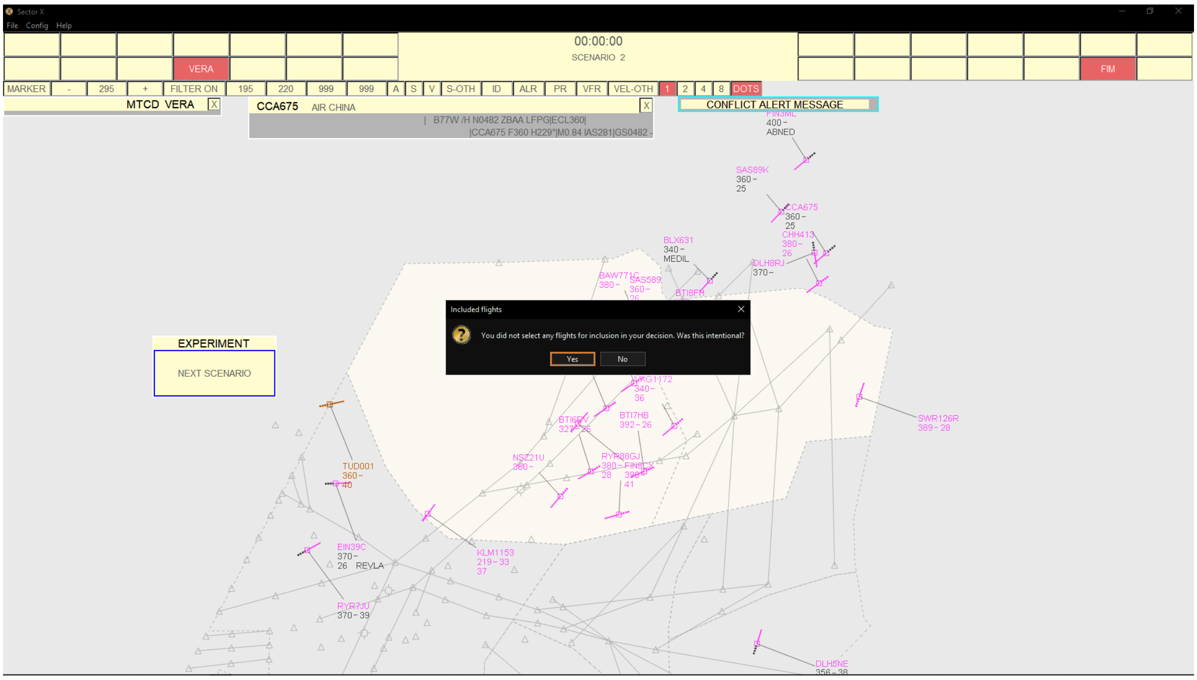


Figure B.5: Screenshot of a Jever scenario with the included flight message

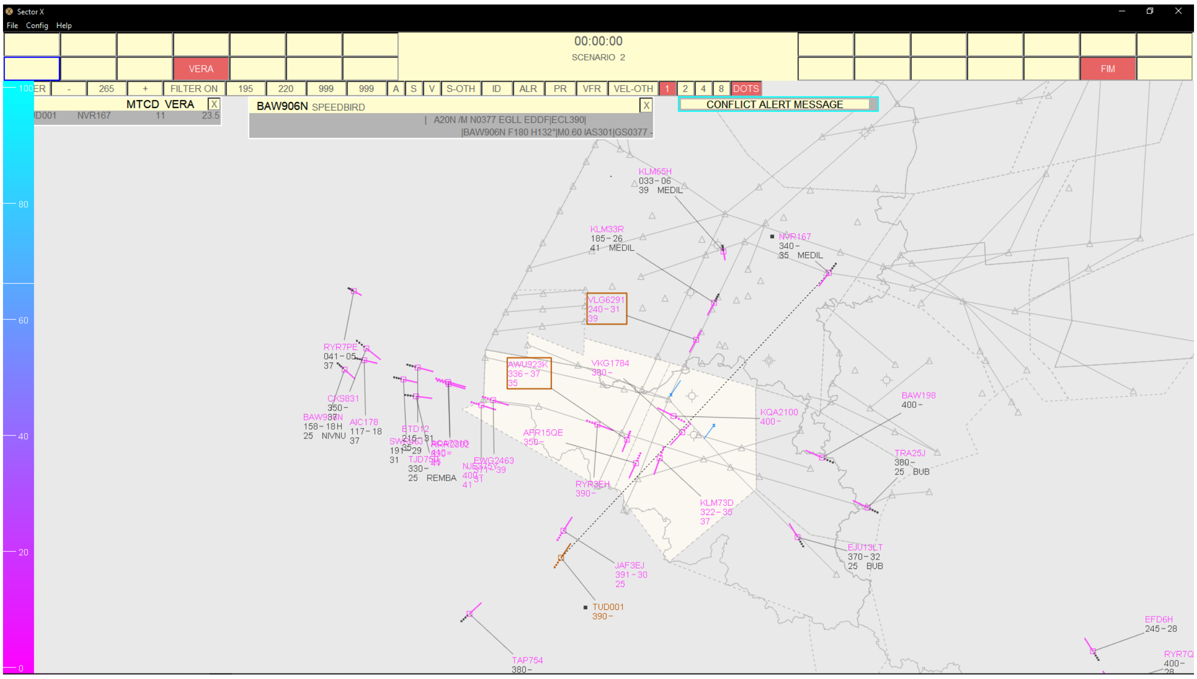


Figure B.6: Screenshot of a Brussel scenario with the complexity rating bar on the left side of the radar screen

C
Questionnaire

Part I – participant information

1. . Which sector group are you licenced for?
    - Brussel
    - DECO
    - Hannover
2. What is your gender?
    - Male
    - Female
    - Other
3. How many years of experience do you have as an ATCo? (textbox)

Part II – scenario review (repeated for nine scenarios)

1. How certain are you of your complexity score
    - Extremely uncertain
    - Moderately uncertain
    - Slightly uncertain
    - Neither certain nor uncertain
    - Slightly certain
    - Moderately certain
    - Extremely certain

Question about the flight of interest:

2. What factors regarding the flight of interest did you consider when giving this complexity score? (Multiple answers possible)
    - Possible loss of separation
    - The solution space available to solve a possible interaction
    - Time left before a possible interaction
    - The XFL of the flight
    - The effect of the flight on the rest of the sector plan
    - The flight was not on a DCT trajectory
    - Other:
3. Did you consider anything else when giving the complexity score? (textbox)

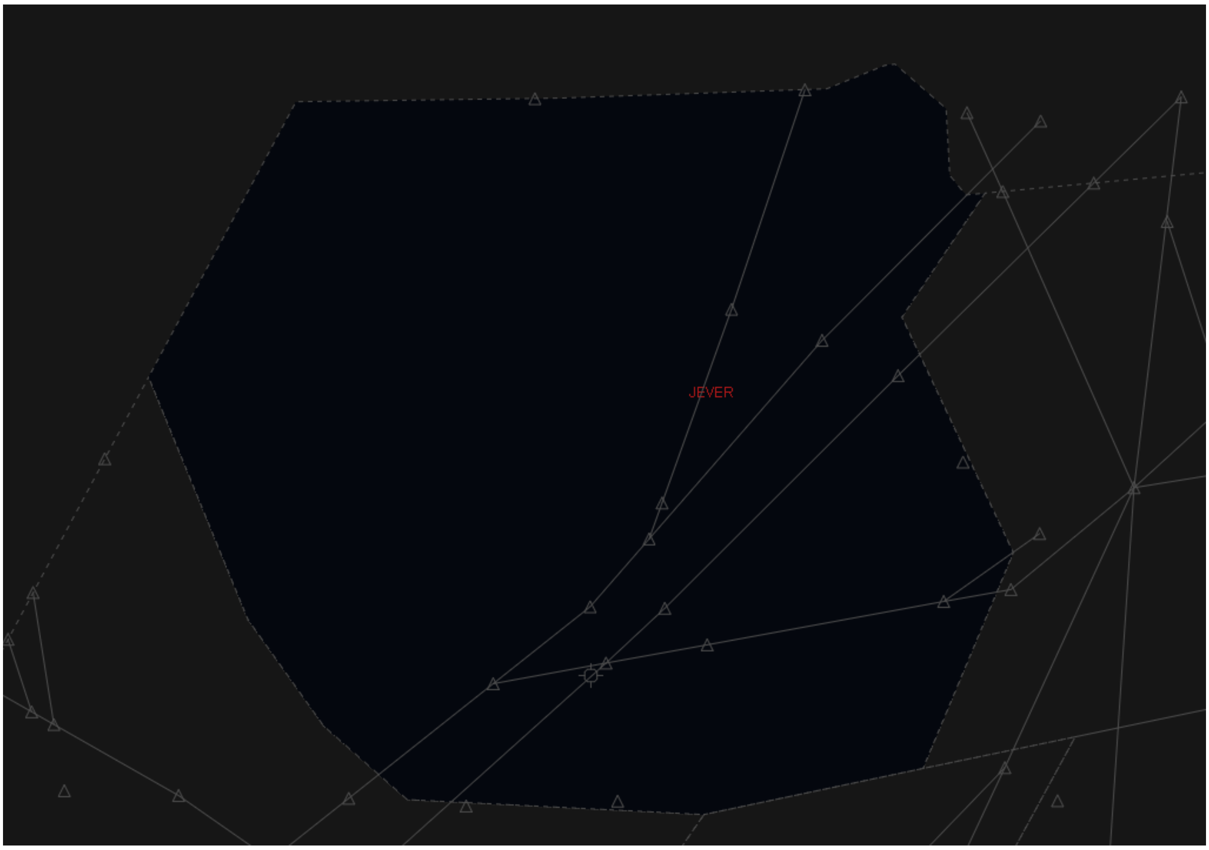
Questions about the included flight:

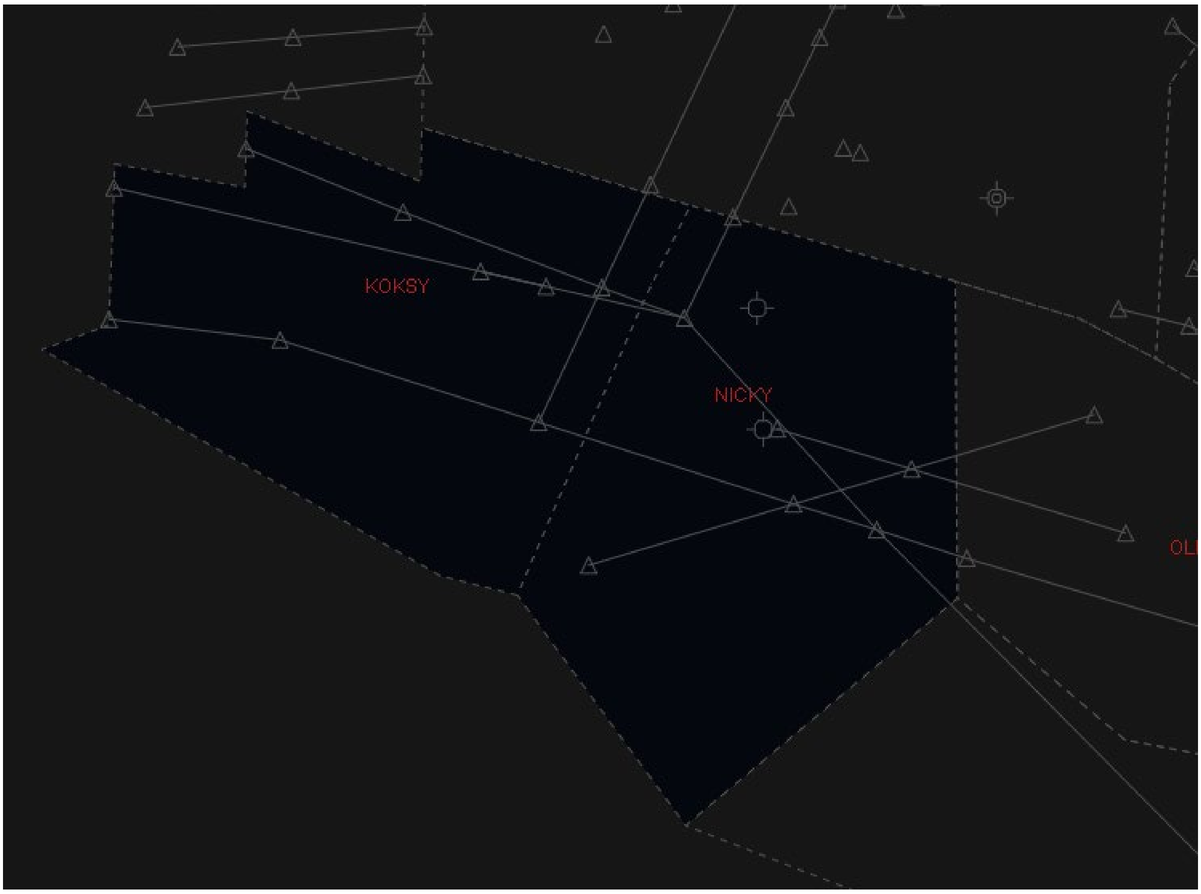
4. Which flight(s) (other than the flight of interest) impacted your complexity score? (Please rank them in order of importance)
5. What factors added to the complexity of this interaction? (Multiple answers possible)
    - Possible loss of separation
    - The solution space available to solve a possible interaction
    - Time left before the interaction
    - The flight level of this flight
    - The fact that the flight was climbing/descending
    - The XFL of this flight
    - Other:
6. Did you consider other factors? If so, what were they? (textbox)

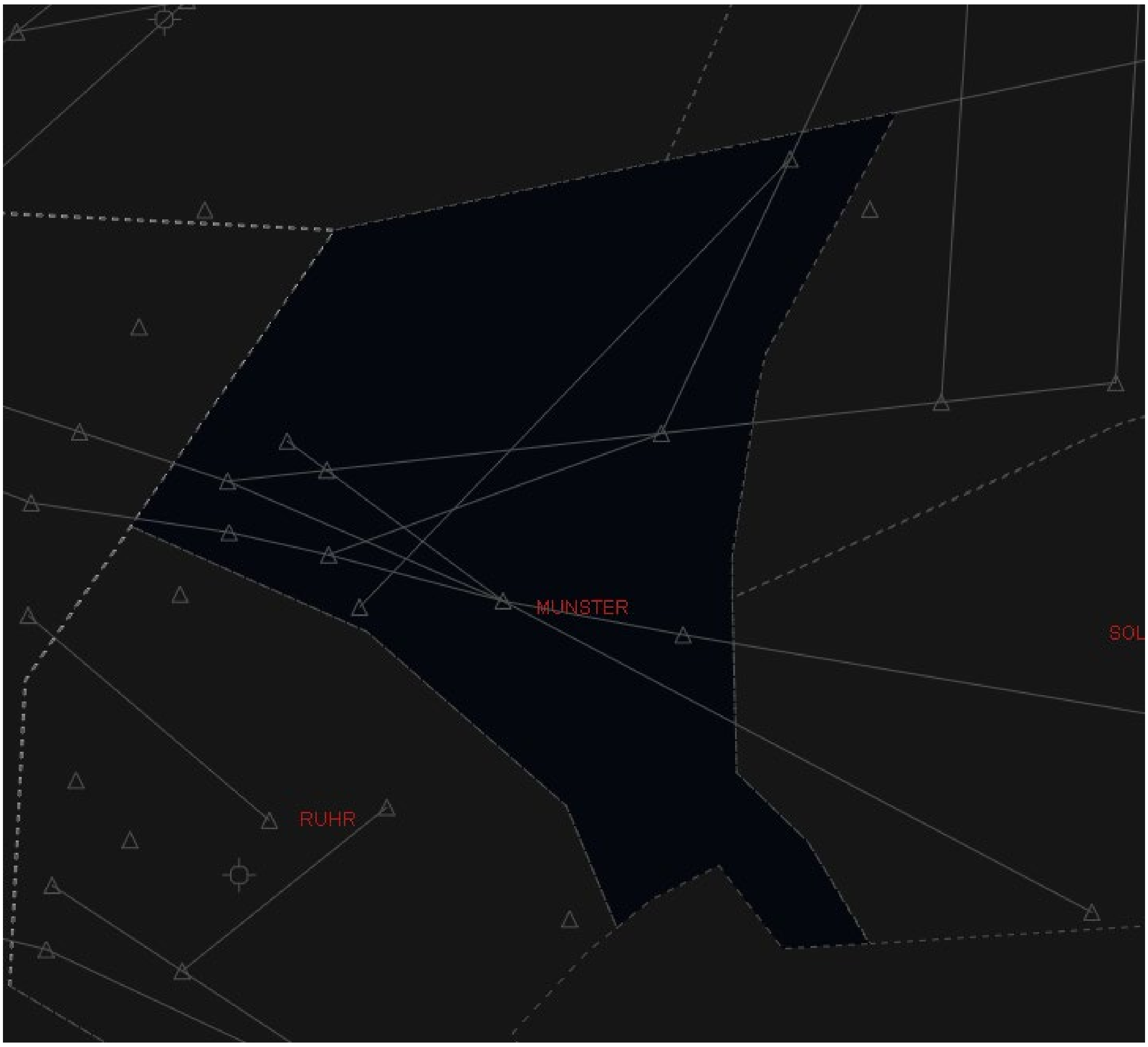
7. Would you be comfortable with the flight of interest being handled by automation?
    - Extremely uncomfortable
    - Somewhat uncomfortable
    - Neither comfortable nor uncomfortable
    - Somewhat comfortable
    - Extremely comfortable
8. What would have to change in the scenario for you to be comfortable with the flight of interest being handled by automation? (textbox)

Part III – questions about the simulation

9. How would you rate the realism of the this experiment for the following aspects?
    a. Traffic scenario
        - Not realistic at all
        - Slightly realistic
        - Moderately realistic
        - Very realistic
        - Extremely realistic
    b. Look and feel of the simulator
        - Not realistic at all
        - Slightly realistic
        - Moderately realistic
        - Very realistic
        - Extremely realistic
    c. Use of tools in the simulator
        - Not realistic at all
        - Slightly realistic
        - Moderately realistic
        - Very realistic
        - Extremely realistic
10. . Can you indicate spots on the sector map where complex situations frequently occur? You can click on the spots on the map. To remove a spot click again.


JEVER


KOKSY
NICKY


MUNSTER
SOL
RUHR

D
Informed Consent Form

Consent form

Complexity of Individual Flights in En-Route Air Traffic Control

Please tick the appropriate boxes

Taking part in the study	**Yes**	**No**
I have read and understood the study information dated 2022/04/XX, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	○	○
I understand that taking part in the study involves having simulation data automatically stored in an anonymous manner when completing the experiment.	○	○
I understand that taking part in the study involves me answering questions to surveys.	○	○
I understand that taking part in the study involves taking notes about the things I do or say.	○	○
Use of the information in the study	**Yes**	**No**
I understand that information I provide will be used for analysis and (scientific) publications on an anonymous basis.	○	○
I understand that personal information collected about me that can identify me, such as my name and email address, will not be shared beyond the study team.	○	○
I agree that my information can be quoted in research outputs on an anonymous basis.	○	○
Future use and reuse of the information by others	**Yes**	**No**
I give permission for the recorded simulation data and answers to surveys that I provide, to be archived in secure folders, so it can be used for future research and learning. All data is stored anonymously. Access is safeguarded and not to be used for commercial use.	○	○

Signatures

________________________ ____________________ _________

Name of participant Signature Date

----- To be completed by researcher -----

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

________________________ ____________________ _________

Name of researcher Signature Date

Amber Stienstra | +31650818931 | A.Stienstra@student.tudelft.nl

E

Complexity Metric

For the design of the scenario the outputs of the metric could be viewed in the design screen. The outputs were updated every time a flight was moved. The complexity viewed showed the following characteristics of the interactions of a selected flight with the other flights on the scenario:

- Min horiz. CPA [NM] = Minimum distance between the two flights in the horizontal plane
- Min vert. CPA [NM] = Minimum distance between the two flights where the vertical separation is less than 1000 ft.
- Angle [deg] = Conflict angle
- Time to CPA [s] = Time remaining before moment where the distance between the two flights is minimal
- Time to LOS [s] = Time remaining before the two flights violate the separation minima

The latter three characteristics are only displayed in case there is less than 1000 ft vertical separation between the two flight anywhere on their predicted trajectories. Figure fig. E.1 shows a screen shot of the complexity viewer, the selected flight (TUD001) is shown in the top bar.

The interaction characteristics are calculated for eight different trajectories:

- No action = Current trajectory no action
- Direct = Flying direct to the sector exit point
- Route @ TFL = Staying on the current route, but going to the sector transfer level
- DCT @ TFL = Flying direct to the sector exit point, and going to the sectort transfer level
- Route + 1000 = Current route 1000 ft climb
- Route - 1000 = Current route 1000 ft descent
- DCT + 1000 = Direct to sector exit and 1000 ft climb
- DCT - 1000 = Direct to sector exit and 1000 ft descent

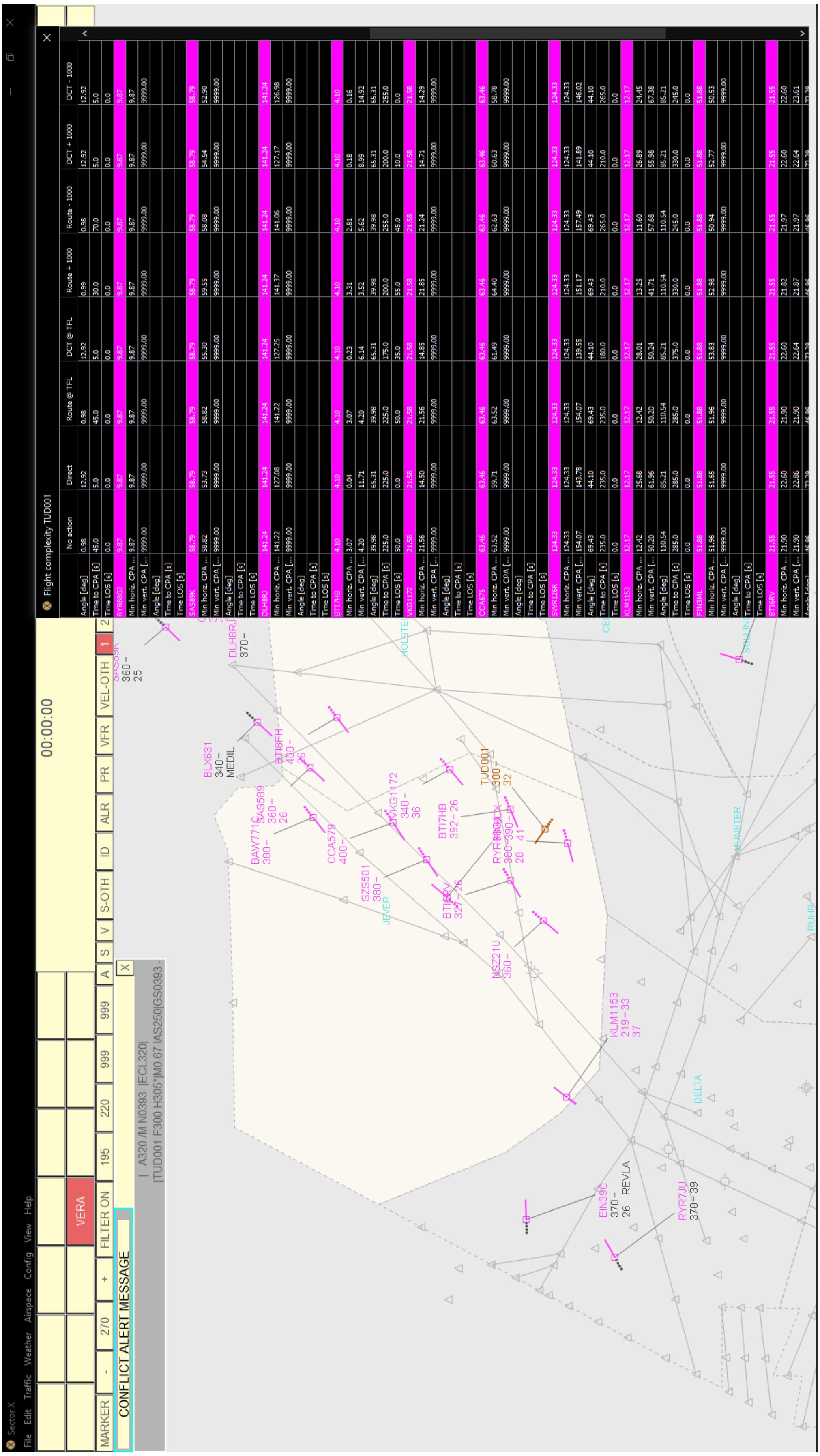


Figure E.1: Screenshot of the complexity metric viewer used for the scenario design

F

Additional Results

F.1. Questionnaire comments

The participants could add additional comments on factors that affect a flight's complexity, apart from the standards list of factors given in the questionnaire. A part of the comments given were related to flight trajectories that did not conform with the standard working procedures for the sector, which added to the reported complexity.
Comment related to a flight's complexity:

- **Brussel Participant 1:** unusual route
- **Brussel Participant 2:** xfl limits the options
- **Brussel Participant 2:** multiple flights have to change fl
- **Brussel Participant 3:** The TUD officially cannot be acted on without release from ext. partner
- **Brussel Participant 4:** speed
- **Jever Participant 1:** time critical because late transfer
- **Jever Participant 1:** The change of level of this flight requires quite some attention, probably some turns of a few aircraft combined with monitoring the climb above other aircraft
- **Jever Participant 3:** vaak meer werken in stappen, 10 min look ahead en klimmen in stappen om onzekerheid weg te nemen, sommige vluchten daardoor geinclude maar eigenlijk pas na de normale look ahead time van belang
- **Munster Participant 2:** rare routes
- **Munster Participant 4:** The aircraft is quite low, with few interactions with other aircraft.
- **Munster Participant 4:** The flight is fairly low in the sector, with fewer departures and arrivals, generally means that there's not so much in the way

For each flight of interest the participants were asked to rate how comfortable they would be with with it being handled by an automated system. The following comments were given in relation to this question.
Comments related to the automation of the flight of interest:

- **Brussel Participant 2:** (I am) more comfortable as the solution can be created with other flights
- **Brussel Participant 4:** Need more experience with automation to be comfortable

- **Brussel Participant 4:** I need to make sure that the actions taken by an automated system to solve this "problem" will not have too much impact on the rest of the traffic i.e. turning an aircraft out of the way to make space for TUD001 will not create a conflict with another flight
- **Jever Participant 2:** automation: comfortable not when conflict is coming, more for small tasks such as hand over and long term planning, depends on how computer would handle
- **Jever Participant 4:** Certainty that the flight would not turn unexpectedly.
- **Munster Participant 1:** If I knew that automation was also considering KLM1918 in the solution.
- **Munster Participant 2:** I can fly around this flight
- **Munster Participant 3:**
- **Munster Participant 1:** If I knew that automation was also considering KLM1918 in the solution.

F.2. Additional graphs paper

Results not printed in the paper.

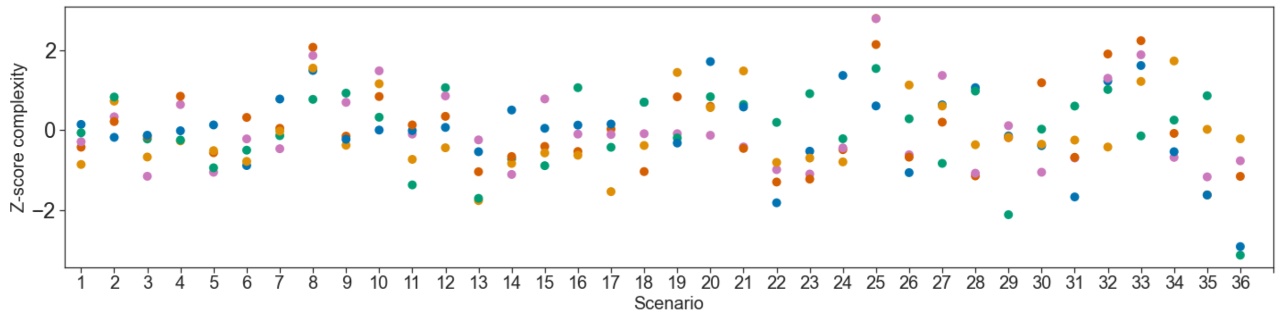


Figure F.1: Z-scores of Munster participants per scenario ordered by spread in complexity score. Each column of scores represents one of the 36 scenarios. The scores for each participants are given in a different colour.

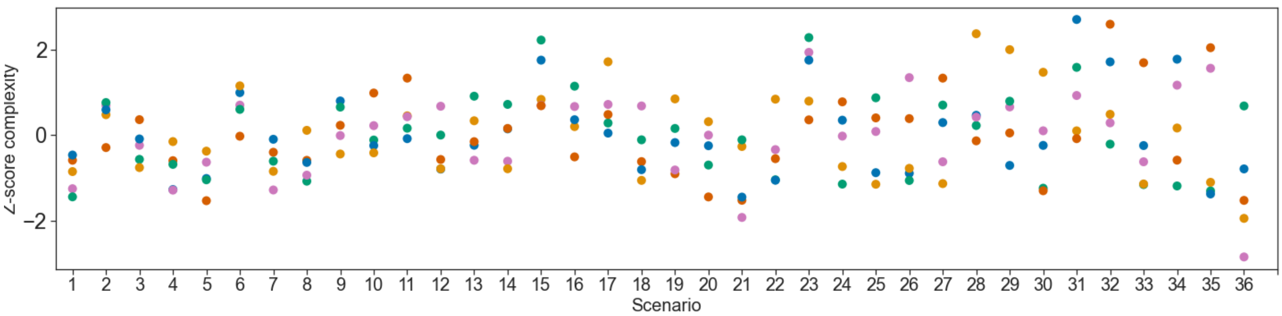


Figure F.2: Z-scores of Jever participants per scenario ordered by spread in complexity score. Each column of scores represents one of the 36 scenarios. The scores for each participants are given in a different colour.

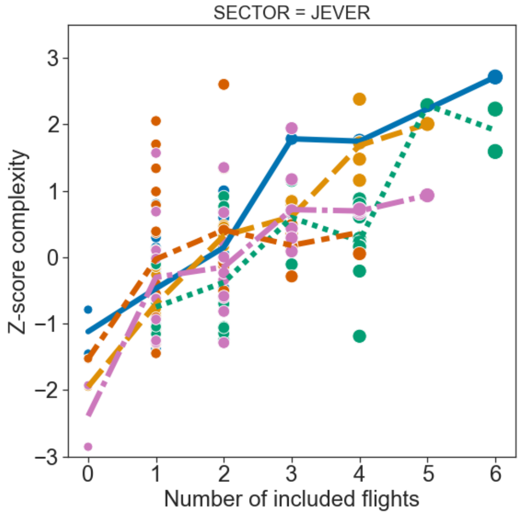


Figure F.3: Number of included flights/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant.

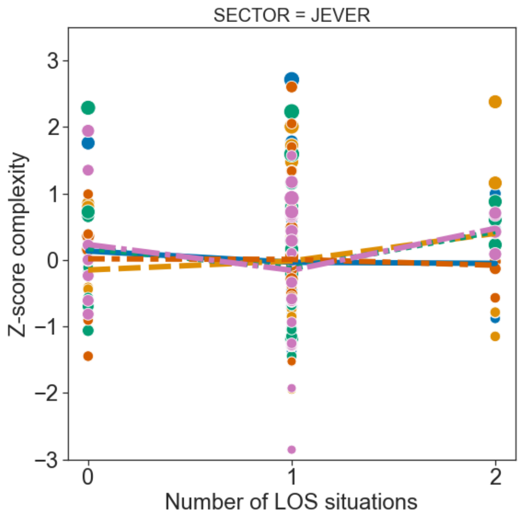


Figure F.4: Number of LOS situations/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participant.

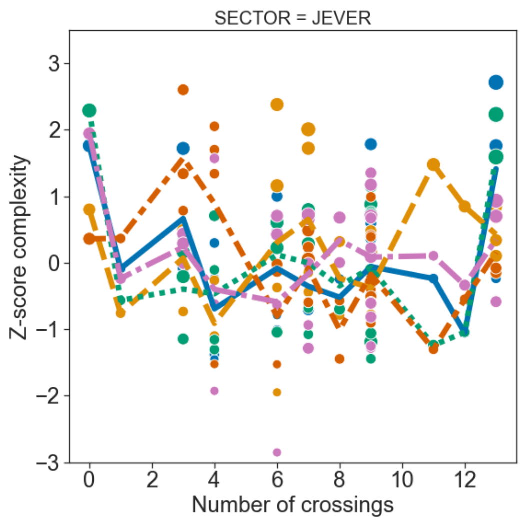


Figure F.5: Number of crossings/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participants if multiple scenarios existed with the same metric outcome.

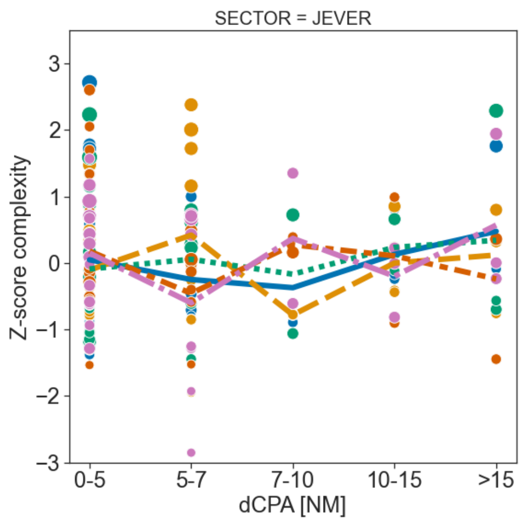


Figure F.6: dCPA/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participants if multiple scenarios existed with the same metric outcome.

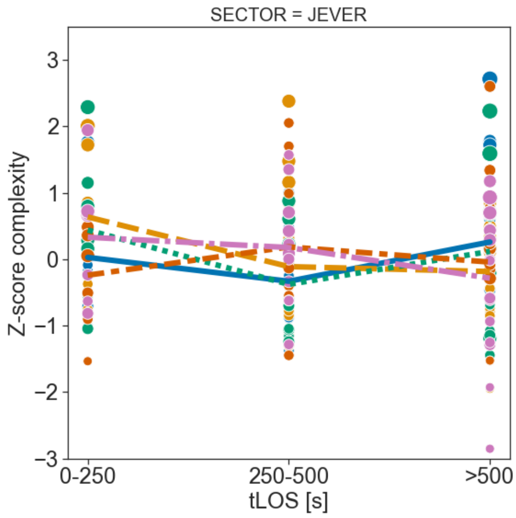


Figure F.7: tLOS/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participants if multiple scenarios existed with the same metric outcome.

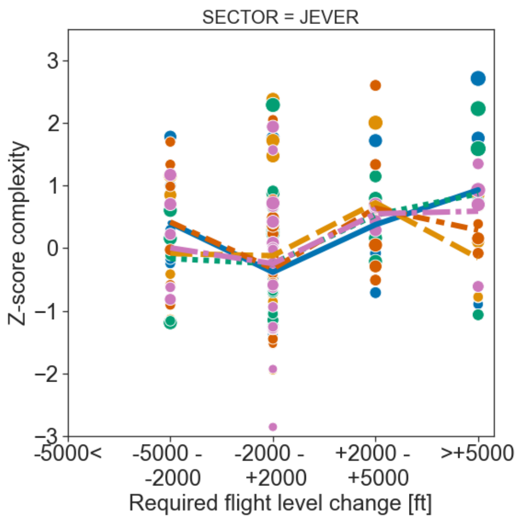


Figure F.8: Required flight level change/Z-score complexity. One colour represents the scores of one participant, the lines are the average score for one participants if multiple scenarios existed with the same metric outcome.

F.3. Further scenario analysis

The use of real traffic samples for the experiment made it difficult to change traffic characteristics in an isolated manner. In an attempt to lessen the effect of the changing variables the metric analysis is repeated here scenarios where there was only one LOS situations. The correlation between the metric outcomes and the reported complexity can be found in Table F.1.

The strength of the correlation between the required flight level change and complexity increases for the Jever and Brussel sectors, and for the number of crossings for the Munster sector. For the number of included flights the correlation strength decreases, and for dCPA and tLOS the correlation is no longer found significant.

Table F.1: Kendall-Tau correlation for all scenarios with one interaction per sector
* significant correlations are in boldface

	Jever		Munster		Brussel	
	Z	p	Z	p	Z	p
Number of crossings	0.10	0.11	**0.26**	**$p < 0.001$**	-0.05	0.52
Number of included flights	**0.54**	**$p < 0.001$**	**0.45**	**$p < 0.001$**	**0.41**	**$p < 0.001$**
tLOS	-0.01	0.88	-0.12	0.15	-0.09	0.21
dCPA	-0.03	0.65	-0.13	0.11	-0.09	0.24
Required FL change	**0.36**	**$p < 0.001$**	-0.04	0.68	**0.40**	**$p < 0.001$**

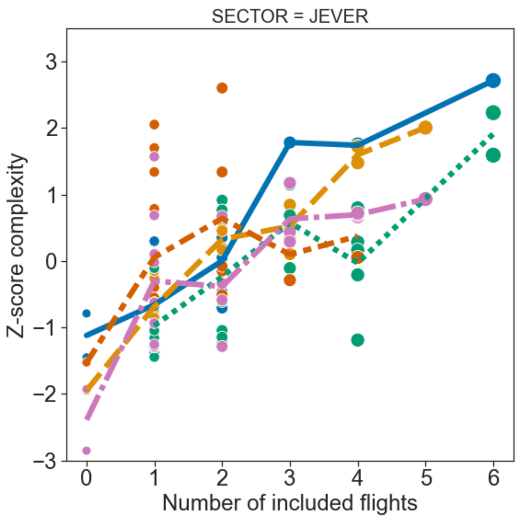


Figure F.9: Number of included flights/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

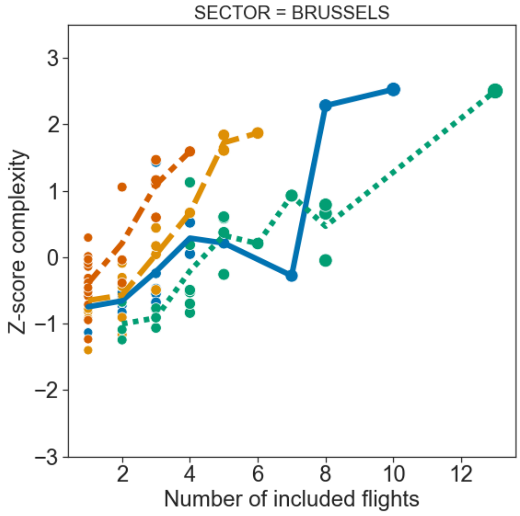


Figure F.10: Number of included flights/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

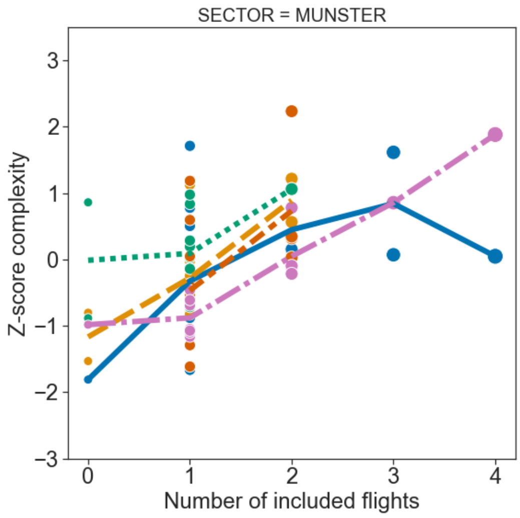


Figure F.11: Number of included flights/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

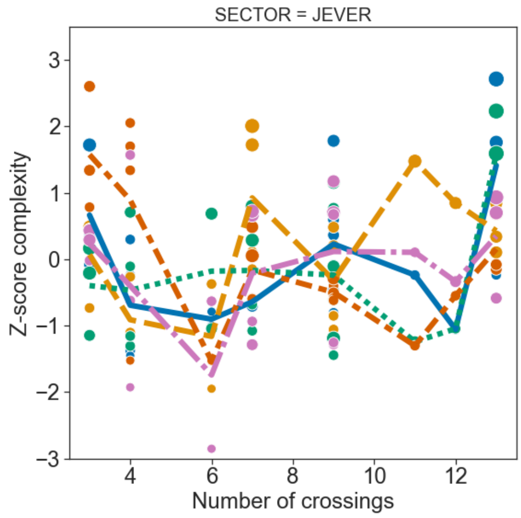


Figure F.12: Number of crossings/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

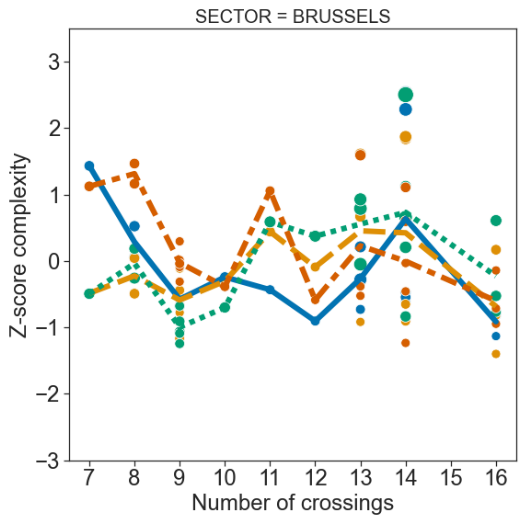


Figure F.13: Number of crossings/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

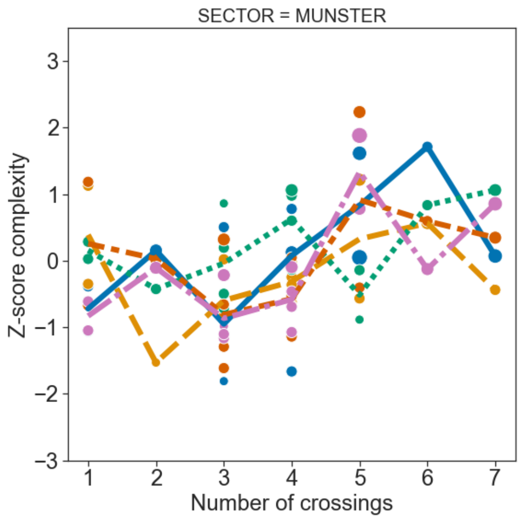


Figure F.14: Number of crossings/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

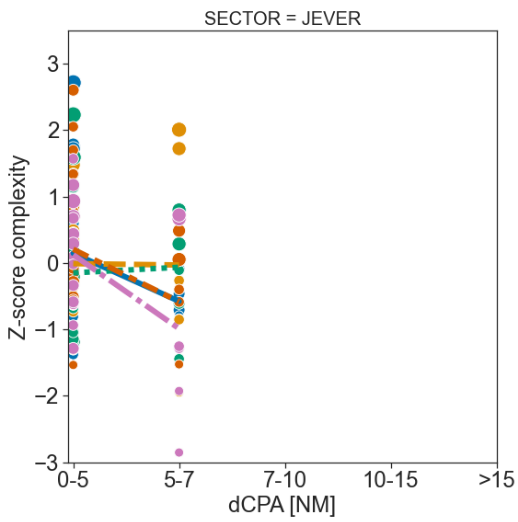


Figure F.15: dCPA/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

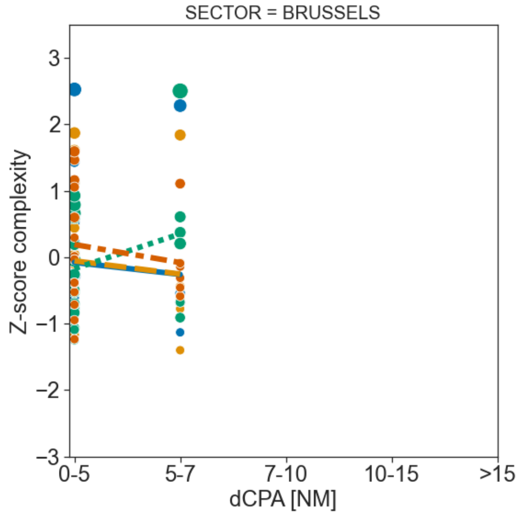


Figure F.16: dCPA/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

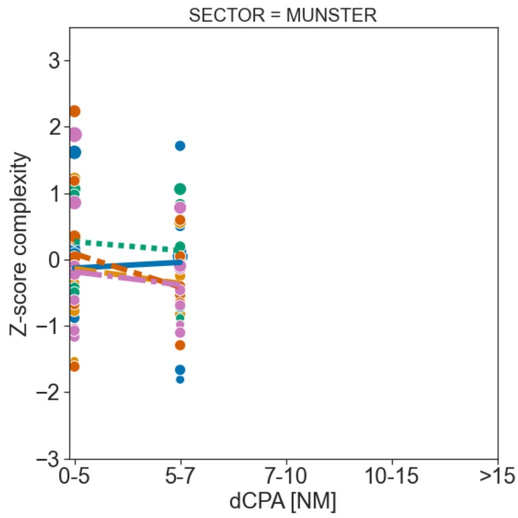


Figure F.17: dCPA/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

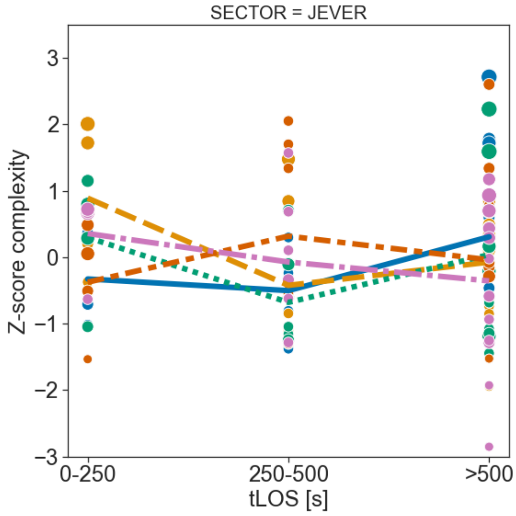


Figure F.18: tLOS/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

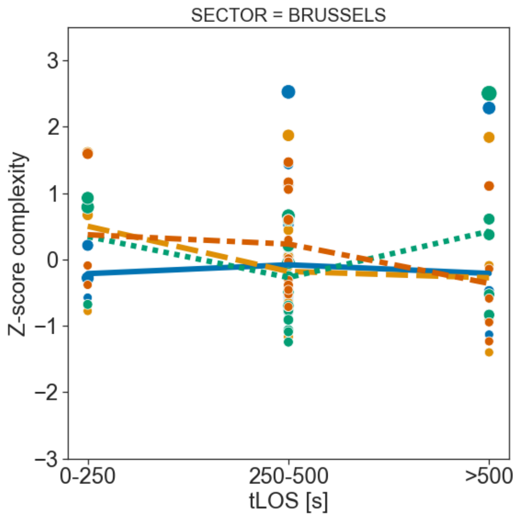


Figure F.19: tLOS/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

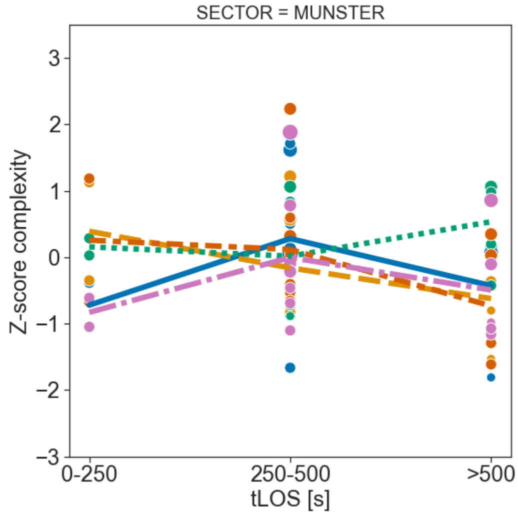


Figure F.20: tLOS/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

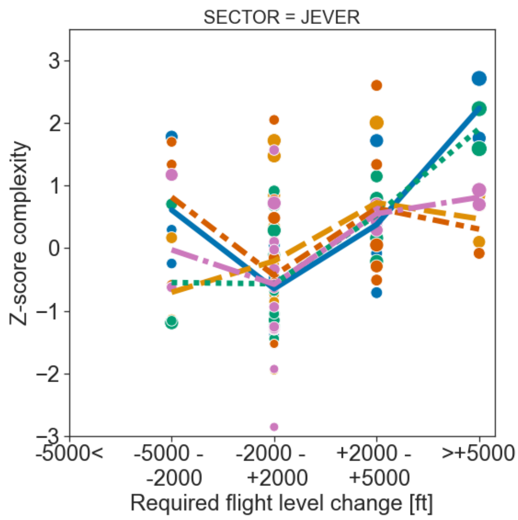


Figure F.21: Required flight level change/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

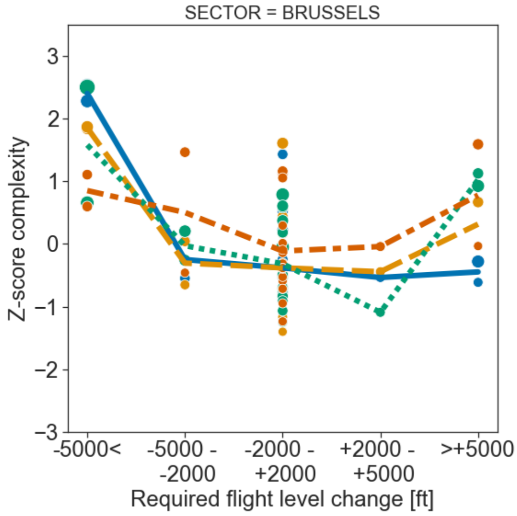


Figure F.22: Required flight level change/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.

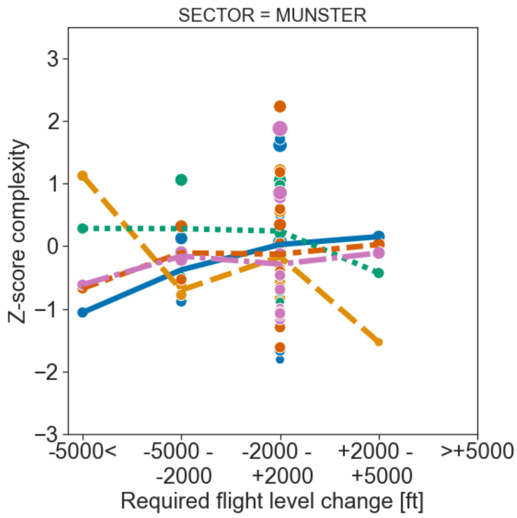


Figure F.23: Required flight level change/Z-score complexity. Scenarios with one interaction. One colour represents the scores of one participant, the lines are the average score for one participant.