Solution Space-Based Approach to Assess Sector Complexity in Air Traffic Control

S. M. B. Abdul Rahman

Solution Space-based Approach to Assess Sector Complexity in Air Traffic Control

Siti Mariam binti Abdul Rahman

Cover design by Joost Ellerbroek

ISBN/EAN: 978-94-6186-255-6

Solution Space-based Approach to Assess Sector Complexity in Air Traffic Control

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op woensdag 5 februari 2014 om 12.30 uur

door

Siti Mariam binti Abdul Rahman

ingenieur luchtvaart en ruimtevaart
geboren te Leicester, Verenigd Koninkrijk

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. M. Mulder

Samenstelling promotiecommissie:

Rector Magnificus	Technische Universiteit Delft, voorzitter
Prof. dr. ir. M. Mulder	Technische Universiteit Delft, promotor
Dr. ir. M. M. van Paassen	Technische Universiteit Delft, copromotor
Prof. dr. ir. J. M. Hoekstra	Technische Universiteit Delft
Prof. dr. R. Curran	Technische Universiteit Delft
Prof. dr. ir. H. A. P. Blom	Technische Universiteit Delft
Dr. ir. J. C. F. de Winter	Technische Universiteit Delft
Dr. B. G. Hilburn	Center for Human Performance Research, CHPR BV
Prof. dr. ir. J. A. Mulder	Technische Universiteit Delft, reservelid

Dr. ir. C. Borst heeft als begeleider in belangrijk mate aan de totstandkoming van het proefschift bijgedragen.

Dit onderzoek is mede mogelijk gemaakt door de financiele ondersteuning van Universiti Teknologi Mara (UiTM), Malaysia.

Summary

Solution Space-based Approach to Assess Sector Complexity in Air Traffic Control

Siti Mariam binti Abdul Rahman

Various methods have been introduced in the past in efforts to optimize airspace sector design and the allocation of air traffic controllers. This is done with the aim to accommodate growth, increase productivity and most importantly to ensure safety of air traffic. To accomplish this, a more comprehensive understanding of human workload, especially that of the controllers involved, is required.

In Air Traffic Control (ATC), there exists a maximum number of aircraft per sector that the Air Traffic Controller (ATCO) is assumed to be capable of controlling simultaneously. The maximum controllable traffic is gathered based on experimentation and subjective assessments of controller workload, which are sector specific. This threshold is not to be exceeded in order to maintain a reasonable and sustainable level of workload. However, a sector complexity metric based on the maximum number of aircraft does not consider the dynamic nature of air traffic, thus limiting the possibility of accommodating the growth of air traffic. Consequently, to better support strategic decisions that need information on ATC workload, we need better measures than just the number of aircraft.

Metrics, for example the Dynamic Density (DD) that use a weighted combination of static and dynamic airspace properties, such as the number of aircraft flying through a sector, the ratio of climbing, cruising and descending aircraft, the horizontal proximity between aircraft et cetera, have been constructed and proposed as a sector complexity measure. The proposed weightings are determined through regression analysis on expert judgement for a particular sector design. As a result, these metrics become highly dependent on sector and operator-centered factors and therefore not uniformly applicable to a wider range of operators and sector designs. A careful calibration would then be needed to tailor the measure to each individual operator and also to the considered sector.

In an effort to find a more objective measure of sector complexity and a predictor of workload, this thesis investigates a constraint-based measure based on the Solution Space Diagram (SSD). In essence, the SSD is a method to observe aircraft restrictions and opportunities to resolve air traffic conflicts in both the speed and heading dimensions. The SSD can be described as the available control area for the controlled aircraft in respect to other observed aircraft within the vicinity. The construction of the SSD is based on the projection of the 'zone of conflict' of the observed aircraft where the key constraint is the 5 NM separation minimum between aircraft. When considering the SSD for any individual aircraft, all neighboring aircraft introduce a 'no-go area' or 'zone of conflict' on the SSD. Intrusion of this zone is called a conflict, or, loss of separation.

Preliminary work conducted with the SSD method indicated that it indeed has the potential to capture the dynamics of taskload and in some cases also predict workload. The goal of this thesis has been to investigate whether the constraint-based SSD method is able to capture the dynamics of air traffic complexity (taskload) in an objective and reliable way, making it useful for future Air Traffic Management (ATM) concepts.

The main hypothesis of this thesis is that the more area that is covered on the solution space, that is, the fewer options the controller has to resolve conflicts, the more difficult the dynamic traffic situation is and therefore the higher the workload experienced by the controller will be.

In this thesis, two main area calculation methods are used. These are the whole unsafe area (A_{whole}) for one particular aircraft and the mean unsafe area (A_{mean}), i.e., the average of the solution space of all aircraft flying in the sector. Both area calculations are used in order to understand the effects of different level of sector complexities on the available solution space. The A_{whole} is calculated using the total area covered within minimum and maximum velocity-heading band (aircraft performance limit) of each individual aircraft. The A_{mean} is gathered using the sum of A_{whole} for all individual aircraft in the sector divided by the total number of aircraft. While A_{whole} represents the constraints that limit each individual aircraft, A_{mean} is a metric that represents the overall sector condition.

The research is designed in such a way that various relevant traffic scenarios or conditions are created by either computer simulations of variable conditions (offline simulations), or by evaluating human performance and workload of control task (online experiments with human in-the-loop). Three different en-route control tasks, namely route merging, conflict identification and resolution and also providing clearances towards assigned waypoints are investigated. This is done in order to evaluate the robustness and versatility of the SSD metric. In the human-in-the-loop experiments, the SSD was used as an offline evaluation method of sector complexity and workload, and metrics based on the SSD were compared in terms of their correlations to controller subjective workload ratings given in the experiment. Each chapter in this thesis presents an attempt taken to further investigate the possibility of measuring sector complexity and predicting workload using the constraint-based approach.

The research begins with the investigation of the effects of various sector complexity constructs on the various SSD area properties. Chapter 3 presents investigation on case studies involving two intercepting aircraft at variable intercept angles, route lengths and speed vectors. Changes in the sector design variables are systematically related to observed changes in the SSD area properties. Aircraft horizontal proximities and intercept angles are two examples of dynamic sector variables that indeed demonstrate notable effects on the SSD. Smaller aircraft proximities result in more area covered on the SSD. When observing incoming aircraft, larger intercept angles result in less area covered on the SSD.

In the human-in-the-loop studies presented in Chapter 4 and 5, it was shown that the SSD has a higher or at least the same level of correlation with the subjective workload ratings given by experimental subjects as compared to the number of aircraft. In an attempt to investigate the possibility of measuring workload of different sector complexity factors, scenarios with varying horizontal proximities, intercept angles, number of streams and traffic density were designed and experimented in a human-in-the-loop experiments. In cases relating to different traffic density, changes in workload as a result of varying sector complexity constructs can be predicted by metrics based on the SSD. Higher traffic density has resulted in a trend of higher workload rating and this is also observed in the SSD area properties.

However, constructing different levels of complexity for various horizontal proximities, different numbers of streams and intercept angles has been a challenging task. This is mainly due to unintentional changes to other factors driving sector complexity, while attempting to impose changes in a particular factor. For example, a change in the number of streams within a sector may also contribute to a change in the aircraft horizontal proximities, when the airspace density would be maintained. As for constructing traffic scenarios with different intercept angles, increasing the intercept angle would mean larger distances between aircraft, when the initial Time To Conflict (TTC) is maintained, or a smaller TTC if the initial distance would be maintained. Thus, the studied effect of one sector complexity factor might be

overshadowed by unintentional changes in other sector complexity factors.

In spite of that, the SSD metric has shown to be a reliable metric, which can still maintain its performance even when investigated using different groups of controllers with varying knowledge and experience on ATC. To more thoroughly investigate the applicability and potential advantages of the SSD metric, in Chapter 6, the metric was compared with a widely accepted complexity metric, Dynamic Density (DD). Based on the investigation, the SSD has shown its capability in assessing the inherent difficulty of ATC situation. The correlation between SSD and workload rating was found to be at the same level or better than number of aircraft and unweighted NASA DD metric. In some cases, the SSD even showed a higher correlation than the weighted NASA DD metric. The SSD metric also has the capability to objectively measure sector complexity where it is found to be less sensitive to inter-controller variability and would also be better transferable across sectors than the weighted NASA DD metric.

Looking at the results of the numerous off-line and real-time human-in-the-loop experiments, the proposed SSD metric shows a promising prospect of being an objective measure of sector complexity and a viable subjective workload predictor. However, these results are based on specific experiment settings, assumptions, and simplifications that were made throughout the research. These simplifications and assumptions (for example by assuming a 2-Dimensional (2D) traffic situation or simplifying the Air Traffic presentation to a basic ATC interface) may have influenced the results, in a way that it may have made the SSD metric appear to be overly promising. However, these simplifications had to be made in order to (1) isolate a single sector complexity construct and (2) eliminate interface demand (e.g., range and radar quality) and other task demand (e.g., standard procedures or radar and Radio Telephony (RT) communication).

As already mentioned, isolating a single sector complexity factor is not an easy task, with sector complexity being an intricate subject. Each sector complexity parameter is inter-related to one another, making it difficult to investigate the effect of a single parameter while not causing another parameter to change. While trying to isolate specific complexity parameters, the investigation of single sector complexity variable (based on scenario of only two converging aircraft) might not deliver the 'same' effect as it would deliver in 'real' situation. However, adding another element by introducing other non-conflicting aircraft in the sector might interfere with the controller's attention from the issue that is being investigated. Thus, a trade-off had to be made between investigating single element of sector complexity variable and presenting a natural traffic condition to the subjects.

Secondly, in the attempt to minimize interface and other unrelated task demand, simplification in the experiment settings and the simulator functions have resulted in the simulator resembling only a portion of the controller's work. The lack of simulator realism might have affected the subjective workload ratings and strategies. Additionally, there were no punishable or detrimental consequences for the controller's actions, so controllers were generally bolder in trying out new strategies. The limited realism of the simulation also affected the sense of danger and stress in controlling air traffic. That is, even if controllers failed to maintain separation, it will only affect their performance during the experiment, but no lives were at stake. The simulator also assumed fast and identical responses to controller commands. This might also changed the controller's usual behavior as it may have triggered an intentionally delayed command to resolve a traffic conflict.

Thus, to prove that the method was found to be the most suited metric in measuring sector complexity, a more extensive research regarding its performance and robustness should be done in the future. More comprehensive research on sector complexity has to be done in order to have a better understanding of sector complexity and controller workload. Also, to keep up with the relevance of the current situation, the extension of the SSD to the third dimension is crucial.

Contents

1 Introduction

1.1 Background

Air Traffic Controllers (ATCOs) are responsible for the supervision of a safe, efficient and orderly flow of air traffic. Current Air Traffic Control (ATC) uses conventional technology (e.g., radar and Radio Telephony (RT) communication) and little automation support exists in supervising air traffic. The 2012 Annual Safety Review report by the European Aviation Safety Agency (EASA) (EASA, 2012) indicates that there has been a steady increase in the number of reported Air Traffic Management (ATM)-related safety occurrences from 2008 and 2012. Here, occurrences are defined as accidents, serious incidents and incidents. In 2012, the category that has the largest proportion of 'serious' and 'major' risk bearing occurrences are incidents which are related to separation minima infringements. This category refers to occurrences in which the defined minimum separation between aircraft has been lost. With the predicted growth of world passenger traffic of 4.7% annually (Airbus, 2012), it is important to investigate the causes of these incidents and explore the possible counter-measures.

Initiatives to design future ATM concepts are being launched in both Europe and the United States, within the framework of the Single European Sky ATM Research (SESAR) (Eurocontrol, 2010) and Next Generation Air Transportation System (NextGen) (FAA, 2011) programs, respectively. In the future ATM concepts developed by SESAR and NextGen, an increased reliance on airborne and ground-based automated support tools is anticipated to increase safety. Concepts such as Free Route Airspace (FRA) will be introduced within the future ATM design, which permit aircraft to fly preferred routes, while performing self-separation, with minimal ATC intervention. With the application of concepts like FRA, a more active 'monitoring' role of ATCO's is anticipated.

Although more aspects of air transportation are being automated over time, the task of supervising air traffic is still performed by human controllers and is therefore limited by human performance constraints (Costa, 1993). Without counter-measures, the rise in projected air traffic (Airbus, 2012) would inevitably result in a further increase in the workload of ATCOs. The latter is often cited as one of the main impediments to the growth of air transport (Janic, 1997, Hilburn, 2004, Koros et al., 2004).

To enable air traffic growth while ensuring the safety of air traffic, we need a better understanding of where the workload comes from. There is one main distinction generally made between task demand load (in this thesis referred to as 'taskload') and mental workload (in this thesis referred to as 'workload'). Taskload refers to the objective demands of a task, whereas workload addresses the subjective demand experienced by the operator in the performance of a task. In the effort to distinguish between taskload and workload, Hilburn & Jorna (2001) defined that system-related factors such as airspace demands, interface demands and other task demands contribute to taskload, while operator-centered factors like skill, strategy, experience and so on determine workload. This is illustrated in Figure 1.1.

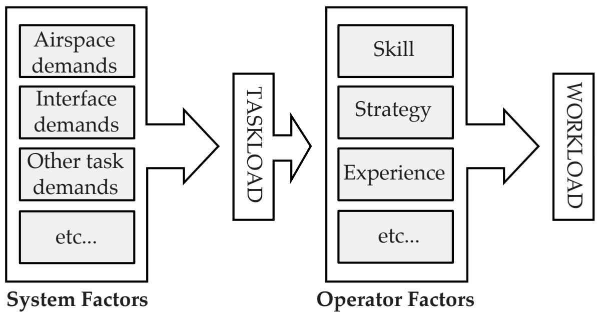


Figure 1.1: Taskload and workload relation (Hilburn & Jorna, 2001).

In literature, sector complexity is a concept introduced to quantify the difficulty and effort required to safely perform the task of air traffic control. Elements of airspace demand, which consists of air traffic patterns and sector characteristics, are the components of sector design, which together define the sector complexity. Airspace complexity depends on both structural and flow characteristics of the airspace (Sridhar et al., 1998). These characteristics represent the static and dynamic aspects to constructing sector complexity, respectively. A good sector design would ensure safety by avoiding high workload for the controller and at the same time promote an efficient flow of air traffic. In order to achieve this dual objective, it is important to find out how the various sector design variables (such as sector shape, the number of routes, the intercept angles of these routes and other constraints) affect the complexity of the task of ATC (taskload) and with that the workload of the individual controller. Studying this relationship forms the main subject of this thesis.

1.2 Problem Statement

Complexity in itself is an ill-defined concept. In order to understand workload in relation to sector complexity, a method to quantify and predict the ATCO workload is needed. Preferably, an objective measure of sector complexity would be available to determine the level of work demand imposed on an 'average' controller, the taskload.

In normal ATC practice, for every sector there is a maximum number of aircraft that the ATCO is assumed to be able to control simultaneously at a reasonable and sustainable level of workload (Majumdar et al., 2004). Whenever demand exceeds sector capacity, three solutions are available; more controllers can be assigned to the sector, a single sector can be divided into two or more sectors, each of which is assigned to its own team of controllers, or aircraft predicted to fly through the sector are deliberately being delayed. However, this concept of limiting the maximum number of aircraft per sector will become less relevant with more complex, future air traffic situations. For instance, in FRA the absence of any route structure will very likely result in a more difficult sector for the same total number of aircraft. Hence, better methods to measure the taskload of controllers, other than simply counting the number of aircraft in a sector, are mandatory.

To include aircraft dynamics behavior, metrics constructed using a weighted combination of scenario properties (such as number of aircraft involved, the ratio of climbing, cruising and descending aircraft and so on), determined through expert judgement and regression analyses were proposed as sector complexity measure (Laudeman et al., 1998, Sridhar et al., 1998, Chatterji & Sridhar, 2001). However, the factor weightings were applicable only in the sector in which they were collected and validated (Hilburn, 2004) and therefore, not uniformly applicable to a wide range of sector designs. Also, the difficulty in using controller complexity ratings or even workload ratings in producing a complexity measure is that it becomes highly subjective, and therefore, careful calibration would be needed to tailor the measure to each individual operator.

Another problem is that (perceived) operator workload is highly dynamic, thereby, it is not only dependent on contextual factors (such as traffic state, weather conditions, sector layout and etc.), but also dependent on the operator's own actions. That is, an operator can influence his own workload by the decisions he makes and be totally unaware of how he actually influenced his own future workload (or task complexity).

A recent study on occurrences of Short Term Conflict Alert (STCA) warnings highlighted that a large number of these alerts do not occur in isolation, but were linked to earlier alerts (Lillo et al., 2009). In fact, over 50% of the STCAs are linked with another STCA and also 23% of the STCAs are multiple STCAs that were involved with other aircraft in a sort of chain (or cascade) process. This indicates that many alerts were caused by ATCO reactions to earlier alerts, i.e., self-induced taskload (or complexity) that is an underexposed element or dimension of perceived workload. In the effort to balance air traffic growth demand and airspace capacity, describing the dynamics of sector complexity is important.

1.3 Metrics for Sector Complexity

1.3.1 Existing Metrics

A number of projects have been performed in the past that explored the use of sector complexity as a workload measure (Laudeman et al., 1998, Sridhar et al., 1998, Chatterji & Sridhar, 2001, Kopardekar & Magyarits, 2003). Measures such as counting the number of aircraft, or Static Density (SD), which uses the number of aircraft per-sector basis (Sridhar et al., 1998, Hilburn, 2004), in many experiments, present the highest correlation with ATCO subjective taskload ratings (Kopardekar & Magyarits, 2002, Masalonis et al., 2003). However, it has significant shortcomings in its ability to accurately measure and predict sector complexity (Chatterji & Sridhar, 2001, Kopardekar & Magyarits, 2002) due to its inability to illustrate sufficiently the dynamics of the behavior of aircraft in the sector. Figure 1.2 shows an example where an ordered parallel traffic flow of nine aircraft will not exhibit the same complexity rating with the same number of aircraft flying various directions. Thus, the SD method alone is unable to represent the maximum number of aircraft that is manageable by a controller.

Another sector complexity measure such as the Dynamic Density (DD) incorporates the dynamic behavior of aircraft in the sector. The DD metric takes into account ''*the collective effort of all factors or variables that contribute to sector-level ATC complexity or difficulty at any point of time*'' (Kopardekar & Magyarits, 2002). However, the calculation of the dynamic density is based on the weights gathered from regression

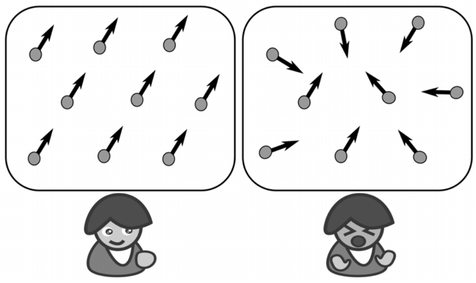


Figure 1.2: Examples of how nine aircraft in a sector can yield completely different complexity measures with different heading angle orientations.

methods on samples of traffic data and comparing them to subjective taskload ratings. As a result, the DD metric represents a complexity measure that incorporates both subjective and objective workload measurements. The method is therefore both sector-dependent and controller-dependent.

Other notions of sector complexity measure using visualization techniques have also been proposed through complexity maps such as the Input-Output (IO) approach by Lee et al. (2009), the Lyapunov Exponents (LE) approach by Puechmorel & Delahaye (2009) and a medium-term multi-sector planning tool called the Tactical Load Smoother (TLS), which was realized during the Programme for Harmonised Air-Traffic Management Research in Eurocontrol (PHARE) project (Whiteley, 1999). However, these complexity maps all have the shortcomings of either being controller dependent (IO approach) or both controller and sector dependent (TLS tool approach) or having a computational challenge (LE approach), which is critical for application to high density airspace.

The long and still ongoing research attempts in this area confirm the importance of exploring ATC sector complexity metrics in understanding its relation to ATCO workload. Clearly, there is a need for an objective metric that can be used to measure taskload in a controller-independent fashion, and also that can be used to compare the complexity of sectors in a quantitative way.

1.3.2 Sector Complexity Measure Proposed in this Thesis

This thesis proposes the use of a constraint-based method, namely the Solution Space Diagram (SSD) as an objective measure of taskload that is independent of controllers and is also capable of being used across sectors. Essentially, the constraint-based method allows for the investigation of the difficulty of a particular air traffic control task, purely based on sector geometric and aircraft kinematic properties. This method describes the constraints that limit the air traffic controller's decisions and actions within the aircraft performance limit.

The method basically works as follows. For any particular ATC situation, the SSD covers all heading and velocity combinations, indicating which velocity vectors offer 'safe solutions' and which velocity vectors lead to an impending conflict with another aircraft. This is represented within the minimum (V_{min}) and maximum (V_{max}) velocity that represent the performance limit of the aircraft. The difficulty of the task is observed based on the examination of the constraint or no-go areas of an aircraft. When considering the SSD in evaluating a sector, each aircraft within the sector introduces a zone of conflict on the SSD of another aircraft. The properties of these conflict zones can be systematically studied to deepen the understanding of the SSD usability in assessing controller workload and sector complexity.

The conflict zones of Van Dam et al. (2004) have been the basis for representing the SSD. It is based on analyzing conflicts between aircraft in the relative velocity plane. Figure 1.3a shows two aircraft, the controlled aircraft (AC_{con}) and the observed aircraft (AC_{obs}). In this diagram, the Protected Zone (PZ) of the observed aircraft is shown as a circle with radius of 5NM (the common separation distance) centered on the observed aircraft. Intrusion of this zone is called a conflict, or, loss of separation. Two tangent lines to the left and right hand sides of the PZ of the observed aircraft are drawn towards the controlled aircraft. The area inside these tangent lines is called the Forbidden Beam Zone (FBZ). If the relative velocity (V_{rel}) falls within the area of the FBZ, future separation violation for the assigned look-ahead velocity vector is foreseen. However, in Figure 1.3a, the V_{rel} falls outside the FBZ, therefore, there will be no PZ intrusion.

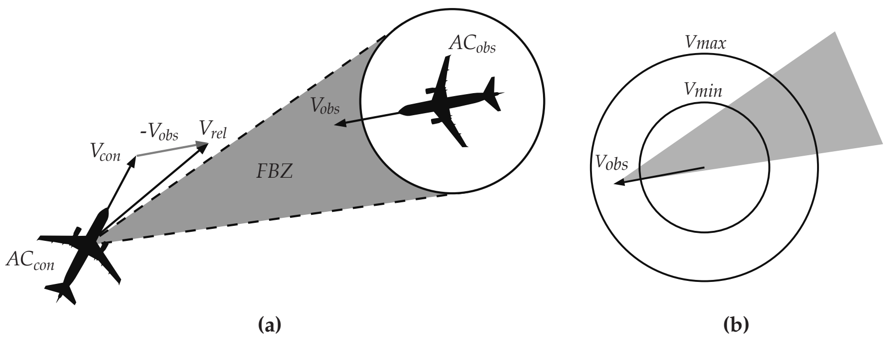


Figure 1.3: Two aircraft condition. (a) Plan view of conflict and the corresponding FBZ. The area within the FBZ represents an instantaneous, complete set of possible conflicting relative velocities. (b) Basic SSD for the AC_{con} by transposing the AC_{obs} velocity vector to the SSD (Adapted from Mercado Velasco et al. (2010)).

From the relative space, the FBZ can be projected to the absolute space by transposing the AC_{obs} velocity vector to the SSD of AC_{con} performance limit as seen in Figure 1.3b. This example serves for a single observed aircraft situation. Additional aircraft would result in more FBZs, thus resulting in fewer options for control. The main assumption is that the fewer options a controller has to control an aircraft, the more complex the task is.

The area covered by the FBZ on the solution space represents the no-go area of an aircraft, defined as the subset of all possible velocity vectors that could lead to future separation violation. The main hypothesis of this thesis is that the more area is covered on the solution space, that is, the fewer options the controller has to resolve conflicts, the more difficult the task and the higher the workload experienced by the controller. Thus, the thesis will investigate whether the non-solution space area of a two-dimensional ATC separation problem can be used to assess the inherent difficulty of ATC situations more accurately and objectively than current metrics.

1.4 Research Objective

In essence, the SSD is a method to observe aircraft restrictions and opportunities to resolve traffic conflicts in both the speed and heading dimensions. Preliminary work conducted with the SSD method indicated that it indeed has the potential to capture the dynamics of taskload and also predict workload (d'Engelbronner et al., 2010). The goal of this thesis is to investigate whether the constraint-based SSD method is able to capture the dynamics of air traffic complexity related taskload in an objective and reliable way, useful for applications of future ATM. Therefore the main research question can be formulated as follows:

''Will a constraint-based approach be able to capture the dynamics of air traffic and sector related taskload in an objective and reliable way, to predict the workload of air traffic controllers?''

1.5 Research Approach

To answer the research question, the main objective is being elaborated from several perspectives:

This research will utilize the two-dimensional SSD method, in its elementary form. It will explore its potential in objectively measuring sector complexity and predicting workload in varying situations. Limiting the investigation based on a two-dimensional environment enables a more focused investigation on the fundamentals underlying the sector complexity construct. Including different flight levels may result in, at this point, an undesired interference with the focus of evaluating individual sector complexity constructs.

In order to assess the reliability of the SSD method, for each simulation or experiment cycle, the theory is revisited and more sector complexity variables are introduced into the investigation (Figure 1.4). While it is important to gain insights into a single factor affecting sector complexity, the reliability and also sensitivity of the metric can be affected by other means. This includes the combination of multiple factors, as well as the creation of scenarios that represent uncertainty in traffic flows.

The research is designed in such a way that various relevant traffic scenarios or conditions are created by either computer simulations of variable conditions, or by evaluating human performance and workload in certain well-defined control tasks (e.g., merging aircraft on a route, identifying and resolving conflict pairs, or giving traffic clearances towards assigned waypoints). By designing various relevant traffic scenarios or condition, we were able to create different levels of air traffic sector complexity constructs.

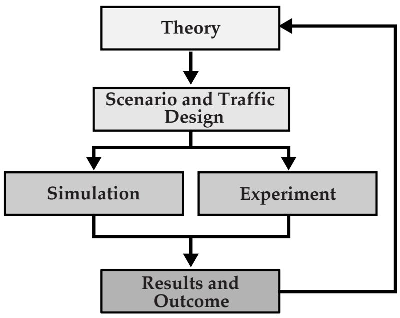


Figure 1.4: Research approach.

The metrics based on the SSD method consider the SSD area percentage measures (both individual and/or average SSD area properties) in order to quantify the level of sector complexity. Conclusions from previous work stated that the area in the SSD that offers solutions (the free area) has a strong inverse correlation with controller workload (Hermes et al., 2009, d'Engelbronner et al., 2010). Thus, this SSD metric will be adopted and further refined in this research to capture both anticipated and unanticipated complexity constructs that are caused by either sector design variables or ATCO-induced changes in the traffic situation.

Figure 1.5 illustrates the relationships between sector complexity, taskload and workload, which are anticipated from the research. As in literature, this thesis defines

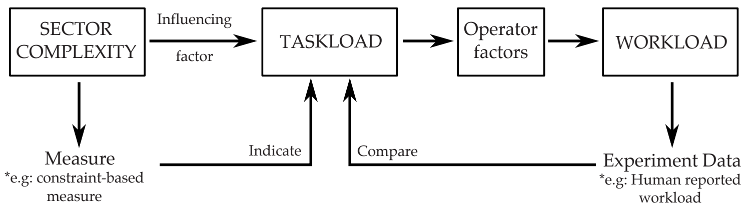


Figure 1.5: Sector complexity, taskload and workload relation.

sector complexity as one of the factors influencing controller taskload. Human workload, on the other hand, is defined as being dependent on operator factors as such that it is addressed based on the demand experienced by individual controller while performing a task. Therefore, in order to predict workload, metrics, which are based on sector complexity constructs (in this thesis, constraint-based SSD measures) are evaluated to indicate the level of taskload. The measured taskload is then compared with the experimental data (reported ATCO subjective workload rating), in order to evaluate whether the proposed measure (SSD) can indeed represent a predictor of controller workload.

1.6 Research Scope and Assumptions

To investigate the effects of sector complexity on taskload, a comprehensive list of complexity variables can be investigated. The list of sector complexity variables that focused on the Area Control Center (ACC) area as developed by Majumdar & Ochieng (2007), combined with the comprehensive list of factors of ATC complexity factors developed by Mogford et al. (1995), is illustrated in Figure 1.6. Based on the listed variables, 12 groups of complexity variables were formulated according to factors relating to weather, traffic, routes, sector and other complexity measures. To reduce the thesis scope, several assumptions have been made beforehand. Only a subset of the complexity variables illustrated in Figure 1.6 will be investigated.

The research scope is defined as follows:

a) **En-route traffic distribution:** Considering the fact that changes in flight levels would occur less during the ACC phase (en-route traffic) as compared to the Approach Control (APP) and Tower Control (TWR) phases, this research focuses on two-dimensional en-route traffic.
b) **Separation task:** An ATCO is not only responsible for the supervision of the efficiency and orderly flow of air traffic, but also for the safety of all traffic in his assigned sector. In current as well as in future ATC environments, even those where 4D trajectory manipulations are envisioned, maintaining safe separation between aircraft will still be the core of the ATCO task. This thesis will therefore

Figure 1.6: Complexity variables.

concentrate on the task of separating aircraft.

The research assumptions are:

a) **2D SSD Limitation:** The scenarios and conditions investigated are formulated based on the current two-dimensional SSD concept. Previous research suggested that ATCOs consider aircraft in pairs, focusing on six attributes of these pairs (in the following order: 1) Flight level, 2) Flight paths, 3) Longitudinal separation, 4) Relative speeds, 5) Direction of flights after reporting points and 6) Lateral separation). The altitude of aircraft pairs in question is always considered first due to its importance in ascertaining conflict likelihood (Leplat & Bisseret, 1966). With this in mind, apart from not including the flight level, the two-dimensional SSD evaluation does include all other five of the six attributes.

b) **Basic ATC interface:** Basic ATC interfaces are used in this thesis to evaluate different sector designs. These may not be similar to those that are conventionally used by ATCOs under current operations. However, the experiment simulators were created with the aim to focus on investigating sector complexity constructs, while minimizing possible effects of interface demand through providing the experimental subjects a direct and elementary representation of airspace.

c) **Limited constraints:** Air traffic controllers involved in the experiments in this thesis are operating in an airspace that is not subjected to constraints other than those related to the particular sector complexity construct under investigation. Examples are weather, standard procedures, radar quality, neighboring sector policy, pilot compliance to instructions and pilot experience.

Taking into consideration the research scope and assumptions, the complete sector complexity construct has been narrowed down to a number of factors, which focus on traffic and route measures as highlighted in black in Figure 1.6. As each sector complexity construct could be inter-related to another, it remains difficult to investigate a single variable while not causing another variable to change. Acknowledging that difficulty, the investigation of sector complexity constructs and their effects on controller workload, will focus on particular main properties.

The main aspects of sector complexity in this research can be divided into either static or dynamic sector design variables. The static variables are: (1) the number of intercept points, (2) the number of routes or streams, (3) the sector shape, (4) the sector volume and (5) clustering of entry and exit points. The dynamic variables are: (1) the intercept angles of traffic routes within a sector, (2) the clustering of intercept points, (3) the geographical location of intercept points, (4) the traffic density, (5) the traffic mix and (6) the traffic proximity. All are highlighted in blue in Figure 1.6.

1.7 Thesis Layout

This thesis consists of seven chapters, and is organized as follows:

Chapter 2: Taskload, Workload, Sector Complexity, and the Solution Space Diagram. This chapter forms the theoretical basis of what constitutes sector complexity, taskload, workload and the SSD. It discusses the findings from previous research on describing airspace sector complexity. It also introduces the solution space-based analysis method adopted in this thesis.

Chapter 3: Solution Space as Sector Complexity Measure. Here the effects of various sector properties on the SSD area will be investigated. The properties of the various metrics, which can be derived from the SSD are systematically studied to deepen the understanding of its use for assessing controller workload. Sector design variables such as traffic horizontal proximity, speed differences, intercept angle, traffic density, and traffic patterns are investigated. In the study of quantitative measurements of sector design parameters it is assumed that a smaller solution space (or equivalently, a denser conflict space) would result in a higher rating for sector complexity, corresponding to a higher level of controller workload.

Chapter 4: Solution Space in Merging Scenarios. Chapter 4 continues the investigation of the SSD method in merging scenarios. There are four sector design variables that were looked into, namely: incoming aircraft proximity, the number of traffic streams, intercept angle, and traffic mix. In addition to that, two groups of subjects, namely 'student' and 'expert' were explored to investigate the reliability of the SSD metric across different types of individuals. Results show that different sector design variables affect controller workload and also SSD properties differently. Despite the fact that both groups performed differently and had different control strategies, the SSD area properties were found to be in a higher correlation in both groups with the controller workload as compared to correlation involving the number of aircraft.

Chapter 5: Solution Space in Conflict Detection Scenarios. This chapter continues the investigation on the use of the SSD as a sector complexity measure, focusing on scenarios regarding an ATCO's ability to detect future conflicts. Two sector design variables are investigated, namely intercept angle and traffic density. The experiment results reveal that higher traffic density leads to higher workload. On the other hand, the intercept angle appears to be a more complicated complexity construct with no common pattern between SSD area properties and workload to be found. The experiment also did not show a clear threshold on SSD area percentage where a controller would start to detect a conflict pair. Thus, while the SSD might be a good measure of sector complexity, it does not represent a trigger for conflict detection.

Chapter 6: Sector Complexity Measures: A Comparison. In this chapter, two different sectors representing two different levels of sector complexity constructs were designed in order to compare the SSD metric with the number of aircraft and Dynamic Density (DD) metric. Based on correlation analyses, it is found that the SSD measure has the highest correlation with controller workload, compared to the unweighted DD metric and the number of aircraft. Construction of a weighted DD metric through regression analysis, demonstrates that, in some cases, the weighted DD metric performed better than the SSD metric. However, an analysis of transfer of the weighted DD metric across sectors and groups of controllers found that the latter metric is indeed more sensitive than the SSD-based metrics.

Chapter 7: Discussion, Conclusions and Recommendations. This chapter summarizes and concludes the analyses presented in the preceding chapters. Conclusions are drawn with regard to the results obtained with the SSD method. Challenges faced during the research and also recommendations for future research on the solution space-based approach are given.

2

Taskload, Workload, Sector Complexity, and the Solution Space Diagram

Developing more advanced human-machine systems for future Air Traffic Management (ATM) concepts requires a deep understanding of what constitutes operator workload and how taskload and sector complexity can affect it. This section introduces taskload, workload and sector complexity and how they are currently assessed and measured. What is missing, however, is a measure that can represent workload independently of sector layout and that is also robust to inter-controller differences. This objective measure of sector complexity can then determine the taskload as imposed on the controller. The chapter introduces the concept of the Solution Space Diagram (SSD), together with the history of its evolution. In this thesis, the SSD-method is proposed as an objective sector complexity measure.

2.1 Introduction

A number of factors affect controller's workload including, but not limited to: airspace complexity, traffic complexity, interface complexity and an individual controller's level of skills and experience. In the effort to balance air traffic growth demand and airspace capacity, describing airspace sector complexity is indeed important. Many efforts have been done in the past to measure and/or predict operator workload using sector complexity (Sridhar et al., 1998, Kopardekar & Magyarits, 2002, Hilburn, 2004, Lee et al., 2009, Puechmorel & Delahaye, 2009). However, most sector complexity metrics that include sector design are calculated according to a set of rules and subjective weightings, rendering them to be dependent of both sector *and* individual controllers.

In this chapter, the main issues on acquiring and utilizing an objective sector complexity measure, in a dynamic environment, are being brought forward. An objective measure is proposed based on the Solution Space Diagram (SSD). This measure is hypothesized to be independent of sector layout and also unbiased by individual differences. Sector complexity, taskload and workload are the three key elements that need a more thorough elaboration and understanding, before evaluating the use of the SSD as an objective sector complexity measure.

2.2 Taskload

Taskload is defined as the objective demand of a task. In distinguishing the difference between taskload and workload, Hilburn & Jorna (2001) defined that system-related factors contribute to taskload, while operator-centered factors determine workload. This is illustrated in Figure 2.1.

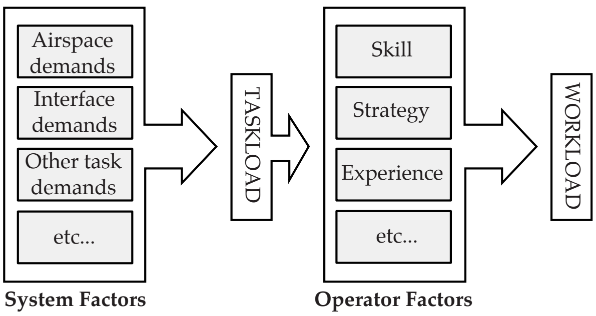


Figure 2.1: Taskload and workload of an ATCO (Hilburn & Jorna, 2001).

It is important to understand the type and quantity of tasks that an Air Traffic

Controller (ATCO) is required to perform at any given time, in order to reveal the influences on the operator's ability to safely and efficiently manage the air traffic. In controlling a complex, dynamic and time-constrained traffic situation, different information resources are used in order to identify and resolve potential conflicts and risky relations between aircraft. These information resources are, amongst others, the radar screen, electronic flight progress strips and Radio Telephone (RT) communications. The ATCO has to perceive, comprehend and anticipate multiple characteristics and flight paths of many aircraft while new incoming aircraft create new traffic relationships to be evaluated. These are examples of taskload-related events imposed on the ATCO, while monitoring and deciding upon the information provided.

2.3 Workload

In order to maintain a safe and expeditious flow of traffic, it is important to optimize the taskload imposed on the ATCO. In the elaboration of taskload and workload by Hilburn & Jorna (2001), the interface demand was included as one of the system factors that contribute to taskload (Figure 2.1). However, according to Mogford et al. (1995), the interface demand can be observed as a mediating factor within the workload definition. Workload according to Mogford et al. (1995) is primarily affected by the situation in the airspace. It is determined by the physical aspects of the sector, for example the sector size, the airway configuration and by factors relating to the movement of air traffic through the airspace such as the number of flights, the number of descending and climbing flights, the number of over flights and lastly by the combination of both sector and traffic characteristics such as the procedures and functions needed (Mogford et al., 1995).

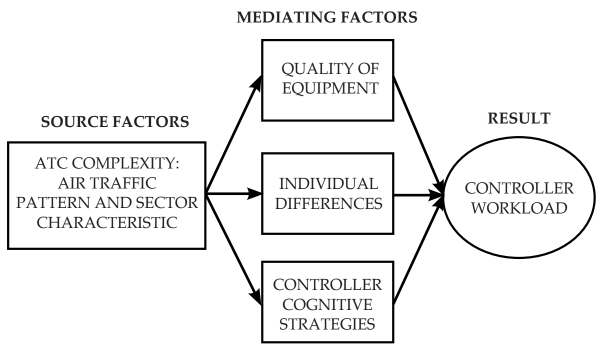


Figure 2.2: Factors affecting ATCO workload (Mogford et al., 1995).

The impact of these primary factors are mediated, however, by secondary factors that include the controller cognitive strategies in processing air traffic information, the quality of equipment, e.g., the human-machine interface and individual differences such as age, proneness to anxiety and experience. The relationship between Air Traffic Control (ATC) complexity and workload, according to Mogford et al. (1995) is shown in Figure 2.2.

Hence, an ATCO is subject to multiple task loads that vary in time. Their performance is influenced by the intensity of the task or demand that they must handle. When coming from a situation of low task demand, higher demands in their tasks will generally yield better performance. However, a demand that becomes too high or too low will lead to performance degradation. Thus, it is important that the demand is acceptable to achieve optimum performance. This will be discussed in more detail in the next section.

2.3.1 Workload Assessment Methods

Workload can be assessed using methods such as performance-based workload assessment, primary and secondary task performance, subjective workload assessment ratings, and lastly physiological measures of workload (O'Donnell & Eggemeier, 1986, Wickens & Hollands, 2000). However, there are issues regarding the sensitivity and diagnosticity of psycho-physiological measures (Scerbo, 2007) and also arguments that cast doubt on finding any direct relation between information load and physiological measures or state estimators (Veltman & Jansen, 2004). Physiological measures are therefore perhaps less suitable in assessing workload and are not further used in this thesis.

Primary task performance and effort are evaluated using two methods: effort-based and performance-based. Effort-based evaluation examples are reaction time and accuracy studies (Farmer et al., 2003) or frequency of individual tasks, such as the number of control actions (Rodgers et al., 1994). Performance-based evaluation examples are either based on the quality of the work, e.g., the number of conflicts that occur or the average deviation from a optimal flight profile.

Secondary task performance is evaluated to determine the amount of 'spare mental capacity' available when the operator is performing the primary task. The evaluation includes counting, calculating, reaction time to auditory and visual stimuli etc. This method measures the decline in performance on the secondary task as a function of the demands on the primary task (Farmer et al., 2003). This is illustrated in Figure 2.3, where variation of demand of the primary task results in variation of available spare capacity. However, there is also a 'willing-to-spend' capacity (Moray, 1977), which is the base level sustainable or acceptable mental load. This willing-to-spend capacity can be different for each individual operator.

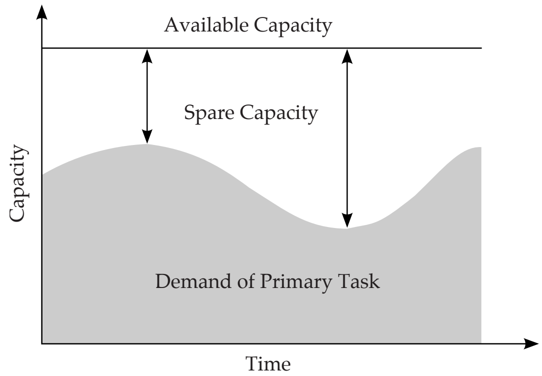


Figure 2.3: Demand of primary task and mental space capacity (Adapted from Farmer et al. (2003)).

The best performance is found around the willing-to-spend capacity, and the interface and tasks should be designed such that low levels of workload, which can lead to boredom, and high levels of workload which can lead to decreased performance, are avoided. This is based on the relationship between workload and performance that is characterized by the inverted U graph as observed by Tulga (1978) and later also reported by Moray (1982). The hypothetical inverted U-shaped curve in Figure 2.4 described an ideal performance-workload relationship. A demand that is either too high or too low for too long, not only constitutes an inefficient application of resources but is also likely to increase the chances of a controller becoming distracted from their primary task, thus leading to performance degradation.

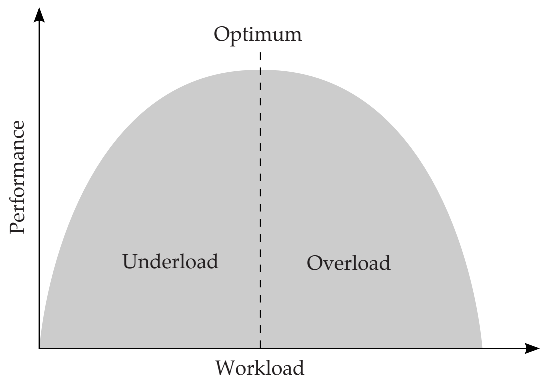


Figure 2.4: Relationship between ATCO workload and performance (Adapted from Tulga (1978)).

The taskload that is imposed onto the ATCO to maintain the productivity, safety and cost factors contribute to workload, where it is agreed to be a subjective and individual response to the task put upon the controller (Mogford et al., 1995, Hilburn & Jorna, 2001). Thus, it is important that the demand is acceptable to achieve optimum performance. To balance demand and performance, an optimum situation where workload and also controller's performance is anticipated.

Another method of assessing workload is through the evaluation of a controller's subjective workload. The resulting workload rating requires the operator to rate the subjective workload at a certain interval. One example of continuous workload rating is Instantaneous Self-Assessment (ISA) that was developed by National Air Traffic Services (NATS) for use in the assessment of ATCO's mental workload during the design of future Air Traffic Management (ATM) systems (Kirwan et al., 2001). The ISA method requires subjects to self-rate their workload during a task, on a scale from 1 to 5 corresponding to, respectively, the lowest workload and the highest workload situation. It is one of the simplest tools with which an estimate of perceived workload can be obtained during real-time simulations or actual tasks (Tattersall & Foord, 1996). The subjective workload can also be measured discontinuously by giving the subjects a post experiment questionnaire such as the NASA Task Load Index (NASA-TLX) (Hart & Saveland, 1988).

In this thesis, a workload rating very similar to the concept of ISA workload rating measure is used and will be compared with the taskload that is estimated from the various airspace or sector complexity metrics, discussed in the next section.

2.4 Sector Complexity Measures

Current-day operations are based on a rather simple, rigidly structured airspace that tends to guide aircraft along fixed corridors and at specific altitudes. The entire path of the aircraft is pre-planned (flight plan) with only minor changes permitted along the route. The control hierarchy is also centralized to ATC, where aircraft can only commence an action upon the approval of clearance requests by the responsible controller.

The area that is controlled by the ATCO is determined as a sector that is defined by fixed boundaries. These sectors may vary in size, depending on the density of air traffic. With a known sector boundary and routing, controllers have a better awareness in terms of areas that need more attention, such as crossing routes or entry points. The ATC task could then be more difficult if aircraft were allowed to fly random routes, because conflicts would arise practically anywhere in the airspace. To overcome the possible excessive taskload demand, this thesis focuses on objectively quantifying sector complexity in order to be able to better predict the effects of sector design variables on taskload, and with that, workload.

Previous studies have indicated that incidents where separation violations occur can happen even when the ATCO's workload is described as moderate (Kinney et al., 1977, Schroeder, 1982). These incidents could have been induced by another factor such as inappropriate sector design. Sector design is one of the key components in the airspace complexity. Airspace complexity depends on both structural and flow characteristics of the airspace (Sridhar et al., 1998). These characteristics represent the static and dynamic aspects to constructing sector complexity, respectively.

A good airspace design would ensure safety by avoiding high workload for the controller and at the same time promoting an efficient flow of traffic within the airspace. In order to have a good airspace design, the ATC complexity variables' impact on controller workload has to be assessed. To achieve this, Majumdar & Ochieng (2007) have listed a number of complexity variables that is subdivided into several major groupings.

In order to have further insights on the effect of sector complexity towards controller's workload, the complexity variables are then ranked based on the controller's rating from 1 (minimum) to 3 (maximum) on the impact of the complexity variables based on their experience. According to the list, variables that impact workload are mostly either traffic, design, weather or the system quality related. However, based on the interviews, it was reported that the ATCOs considered the effect of the combination of complexity variables that increased their workload (Majumdar & Ochieng, 2007).

Table 2.1 lists examples of possible combinations of up to five levels of complexity variables and how they would affect workload. It is observed that with more combinations of complexity variables, a greater level of complexity can be found for a sector.

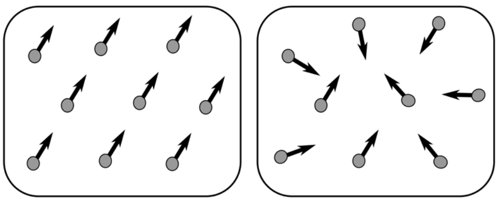
Table 2.1: Five levels of complexity variables (Adapted from Majumdar & Ochieng (2007)).

Level	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5
Level 1	Mix of descends and ascends				
Level 2	Mix of descends and ascends	Long route length			
Level 3	Mix of descends and ascends	Long route length	Route close to the sector boundary		
Level 4	Mix of descends and ascends	Long route length	Route close to the sector boundary	Crossing points	
Level 5	Mix of descends and ascends	Long route length	Route close to the sector boundary	Crossing points	Angle of crossing

In order to understand workload in relation to sector complexity, a method to quantify ATCO workload is needed. One example is by using the sector complexity as an objective measurement indicator. In this thesis, the Solution Space Diagram (SSD) method is investigated as a possible sector complexity measure. Here, our main assumption is that the more solutions an operator has to resolve a traffic conflict, the lower the complexity of the situation and with that, the lower the workload. Before we introduce the SSD method, this section first introduces some sector complexity measures developed in previous researches.

2.4.1 Static Density (SD)

Static Density is one of the measures that are commonly used to obtain an instantaneous indication of sector complexity. It has been the most cited, studied, and evaluated in terms of its influence on workload. In current practice, the complexity of air traffic is generally based on the Static Density (SD) where it is measured based on the number of aircraft per-sector basis (Sridhar et al., 1998, Hilburn, 2004). In many experiments, of all the individual sector characteristics, aircraft SD shows the highest correlation with ATCO subjective workload ratings (Kopardekar & Magyarits, 2002, Masalonis et al., 2003).



(a) (b)

Figure 2.5: Examples of how the same number of aircraft within a sector can yield completely different complexity measures. (a) Nine aircraft with the same heading angle. (b) Nine aircraft with varying heading angles.

The sector throughput depends on the available capacity that a sector has. If the number of incoming aircraft exceeds the available capacity, delays will occur. However, SD has significant shortcomings in its ability to accurately measure and predict sector level complexity (Chatterji & Sridhar, 2001, Kopardekar & Magyarits, 2002). The method has shown to be unable to sufficiently capture the dynamic and kinematic behavior of aircraft in the sector. Figure 2.5 shows an example of different level of sector complexity that could be produced by the same number of aircraft, but with different fixed routing.

2.4.2 Dynamic Density (DD)

Another measurement of sector complexity is Dynamic Density (DD), which is able to include the dynamic behavior of aircraft in the sector. DD is defined as the collective effort of all factors or variables that contribute to sector-level ATC complexity or difficulty at any point of time (Kopardekar & Magyarits, 2002). Research on DD by Laudeman et al. (1998), Sridhar et al. (1998) and Chatterji & Sridhar (2001) indicated a number of dynamic variables (DV) for calculating the DD and each factor is given a subjective weight (W). The DD is a summation of these variables and its corresponding subjective weight:

$$DD = \sum W_i \times DV_i$$

There are several variables selected for inclusion in the definition of the DD function for a sector, such as the traffic density, the number of aircraft with certain heading, speed and altitude change and the number of aircraft within certain lateral and Euclidean distance between each other. A list of complete complexity factors is provided in the literature review by Hilburn (2004). These characteristics were gathered through interviews with qualified ATCOs.

The calculation of the dynamic density is based on the weights gathered from regression methods on samples of traffic data and comparing them to subjective workload ratings. As a result, the DD metric represents a complexity measure that incorporates both subjective and objective workload measurements. It is therefore a controller-dependent method. The assignment of weights based on regression methods also means that the complexity analysis can only be performed on scenarios that differ slightly from the baseline scenario for which the weights were calculated.

To enable the use of DD in varying sectors, the computed weights need to be re-estimated and re-validated for each sector. Therefore the metric is not generally applicable to just any situation (Hermes et al., 2009), rendering the method to be not only controller dependent but also sector dependent. From an operational viewpoint, having too many complexity factors to analyze makes it difficult for decision makers to understand which particular complexity factor is responsible for a high workload situation (Masalonis et al., 2003).

2.4.3 Input-Output (IO)

Lee et al. (2009) propose another sector complexity measure, the complexity map. It is suggested that this method is able to assist air traffic flow managers to identify problematic elements of the sector boundary. The complexity map is constructed based on the control effort or activity required to resolve any arising conflict. If no conflict emerges among the aircraft inside the sector when another aircraft enters the sector, the control activity is zero. When many aircraft need to be given new heading or speed commands due to an incoming aircraft, the control activity is high. Hence, the overall amount of corrective actions needed to recover to a conflict-free condition is taken as a measure of the air traffic complexity. The map is constructed as a function of the position and direction of the incoming aircraft.

There are three different measures of the control effort proposed, with each using different types of control activity measures namely, the sum of the total heading changes over all aircraft inside the sector, the number of heading changes for all aircraft, and the sum of the heading changes due to secondary conflicts. Based on the measured control activity, numerical simulations were performed to illustrate this method of describing airspace complexity (Lee et al., 2009). However, as different control activity measures and different solvers could be used, the choice of the conflict solver has a large impact on complexity evaluation. By using a specific centralized conflict solver in place of an ATCO to safely accommodate a fictitious additional aircraft, this method is in fact also a controller-dependent method (Prandini et al., 2011).

2.4.4 Lyapunov Exponents (LE)

The Lyapunov Exponents (LE) map by Puechmorel & Delahaye (2009) details a dynamical systems modeling of trajectories. The calculation of LE yields a map as a function of aircraft position over the considered airspace area. It identifies places where the relative distances between aircraft do not change with time and the ones where such distances change a lot. Both conditions are interpreted with low and high real LE values, respectively. The larger the positive LE values, the higher the rate at which one loses the ability to predict the system behavior (Prandini et al., 2011), thus the more complex the situation.

This method was found to be both controller and sector independent. However, it comes with one main challenge, which is that it is computationally intensive, hampering its application to a distributed ATM framework. The computational aspects were improved within the frame of the project, but still remain critical for application to high density airspace (Blom & Klompstra, 2011). The method is also seen as not being able to reflect the effect of the level of complexity solely through 'hotspots', or particular locations in the sector with high complexity towards human capacity.

2.4.5 Tactical Load Smoother (TLS)

Another concept of sector complexity measure was developed during the Programme for Harmonised Air-Traffic Management Research in Eurocontrol (PHARE) project. It was developed and used in the medium term multi-sector planning tool called Tactical Load Smoother (TLS) (Whiteley, 1999). The complexity map is calculated as a function of the aircraft vertical evolution, speed, factor for conflict between two or more aircraft, compatibility between an aircraft route and its flight level as well as the distance of the problem from the boundary of the sector. These elements are combined according to a set of rules and weightings, which have been derived both analytically, and empirically using input from experienced air traffic controllers (Meckiff et al., 1998). The complexity map is represented as a heat contour map with hotspots, indicating potential high complexity region in red whereas slightly less complex regions are shown in orange. The objective of the complexity map is to ensure that the controller can see exactly where difficult situations may occur, between which aircraft in the future and from there they can formulate a resolution strategy.

However, the derivation of the TLS complexity map, which includes sector design and also calculated according to a set of rules and weightings, has made the method to be dependent of both sector and individual controller behavior. Again, similar as with the DD method, based on an operational viewpoint, having too many complexity factors to analyze makes it difficult for decision makers to understand which particular complexity factor is responsible for a high workload situation (Masalonis et al., 2003).

2.4.6 Solution Space-based Method

This thesis will investigate a novel sector complexity measure and workload predictor, computed through a constraint-based method, the Solution Space Diagram (SSD) approach. Initial work by Van Dam et al. (2004) has introduced the application of a Solution Space Display in aircraft separation problems from the pilots' perspective. Hermes et al. (2009), d'Engelbronner et al. (2010) and Mercado Velasco et al. (2010) then continued the idea of using the Solution Space in aircraft separation from the perspective of the ATCO. A high correlation was shown to exist between derived metrics from the Solution Space, and the subjective workload reported by ATCOs (Hermes et al., 2009, d'Engelbronner et al., 2010). Figure 2.6 illustrates the evolution of the SSD.

Figure 2.7 illustrates a scenario with eight aircraft and the controlled aircraft corre-

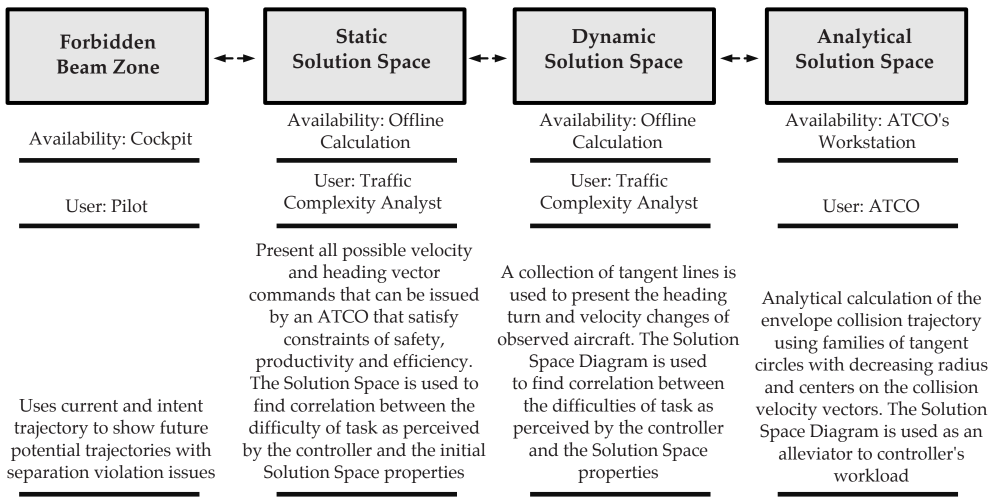


Figure 2.6: Evolution of the SSD.

sponding SSD. The collection of beams (Forbidden Beam Zone (FBZ)) represented in Figure 2.7b illustrates the combination of velocity and headings within the minimum (V_{min}) and maximum (V_{max}) performance limit of the controlled aircraft that will lead to future separation violation with other nearby aircraft in the sector. Different aircraft will have different SSDs, depending on the position, speed and heading of the observed aircraft present in the vicinity.

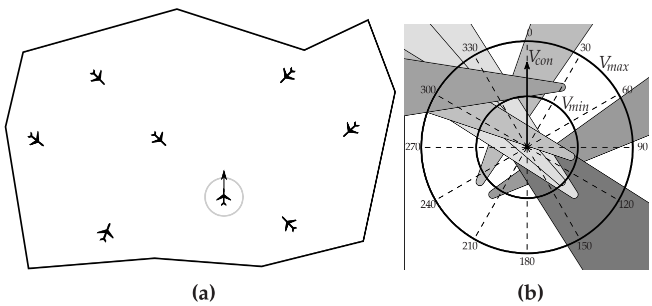


Figure 2.7: SSD example in multi-aircraft scenario. (a) Sector example. (b) The SSD of circled aircraft.

This thesis investigates whether the non-solution space area of a two-dimensional ATC separation problem can be used to assess the inherent difficulty of ATC situations more accurately and objectively than current metrics. The metrics based on the SSD method consider the SSD area percentage measures (both individual and/or average SSD area properties) in order to quantify the level of sector complexity. Figure 2.8a illustrates an example of the SSD of an aircraft, with seven other aircraft within the area. The unsafe area within minimum and maximum velocity-heading band of the respective aircraft (such as in Figure 2.8b) is referred in this thesis as A_{whole}. It defines all possible velocity vectors for the controlled aircraft that could lead to future separation violation.

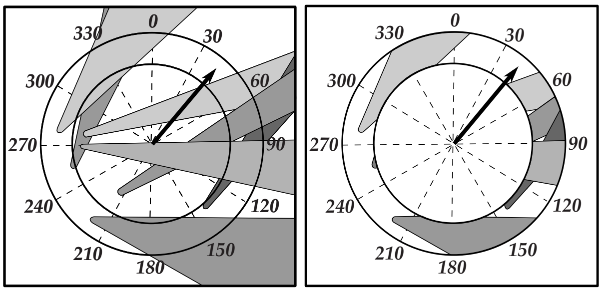


Figure 2.8: The example of SSD unsafe area. (a) SSD with multiple no-go beams. (b) The unsafe area (A_{whole}).

The main hypothesis of this thesis is that, the more area covered on the solution space, that is, the fewer options the controller has to resolve conflicts, the higher the workload experienced by the controller. The extent to which this metric correlates with the self-reported task difficulty of ATCOs is then analyzed through results gathered from human-in-the-loop experiments with various scenarios of varying task difficulty.

The simulation ran four times faster than real-time, similar to what was done in previous research (Hermes et al., 2009, d'Engelbronner et al., 2010, Mercado Velasco et al., 2010). The rationale behind this was to create more variability in traffic situations (and thus workload) within relatively short experimental scenarios. Based on previous research into the assessment of the solution space properties, it is concluded that in a four times fast-time simulation, operators would try to plan ahead the development of events approximately 10 minutes in real-time (d'Engelbronner et al., 2010). With this in mind, the solution space was calculated to provide information relevant to trajectories of known flight plan 10 minutes ahead of real-time. With this setting, the area properties were calculated.

The FBZ represents a collection of relative velocities that will result in a separation violation between aircraft. Each point inside the FBZ corresponds to a certain relative velocity, a certain distance between the aircraft, and also the time at which separation will be lost. The tip of the FBZ represents zero relative velocity and thus corresponds to a time to a loss of separation at infinity. This occurs when the aircraft are flying parallel and thus the separation violation is postponed indefinitely. The further away the relative velocity lies from the tip, the lower the time to a loss of separation will become.

Now given the fact that each part of the FBZ represents the observed aircraft state at a certain moment in time, a cut-off point in time can be determined, enabling the calculation of the SSD that is based on conflicts at less than 10 minutes ahead. The assigned time limit was based on the operator's planned ahead time as mentioned by d'Engelbronner et al. (2010). To demonstrate the effect of cutting off the FBZ, Figure 2.9 shows examples of the SSD at three different cut-off times, namely at infinity (without any cut-off point), a 10 minutes real-time cut-off and a 2 minutes real-time cut-off.

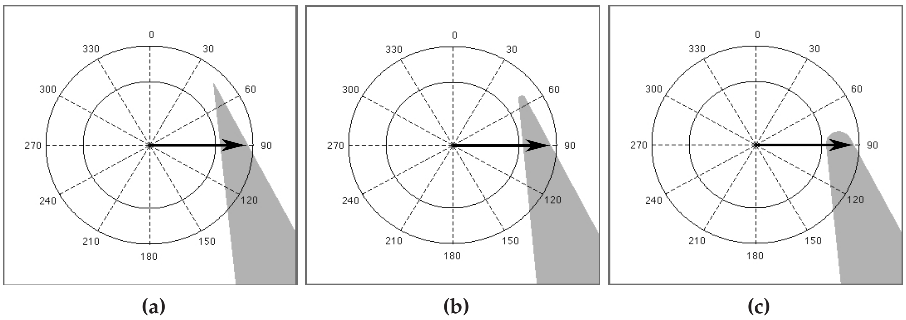


Figure 2.9: Cut-off time in the SSD. (a) Infinity. (b) 10 minutes cut-off. (c) 2 minutes cut-off.

2.5 Requirements for an Objective Sector Complexity Measure

Different studies on sector complexity have been presented in the previous section. Each method has its advantages and disadvantages. In this thesis, we want to investigate and develop an objective measure for sector complexity, independent of sector layout, individual ATCO differences and with good predictive capabilities towards current and future scenarios. To be able to function as a sector complexity measure, a number of criteria (also mentioned in research by Prandini et al. (2011)) are important to consider. The sector complexity measure should be: (a) independent of sector layout, (b) independent of controller differences, (c) able to captures traffic dynamic behavior and (d) able to capture future condition. These are elaborated in the following four paragraphs.

The **sector layout** is an important criterion in currently used complexity measures. Evaluation of the complexity of sectors with a different layout, using e.g., the capacity threshold of a baseline sector may result in a inaccurate sector capacity. Dependency on sector layout also results in rigidity in the sector design. Neither sector complexity measures that rely on the maximum controllable aircraft within a sector, such as used in the SD measure, nor metrics that depend on the workload evaluation based on sector characteristics, such as the DD and TLS complexity map measures, can be considered to be fully independent of sector layout.

Sector complexity is an entity **independent of a controller's individual performance**. A complexity metric that incorporates the ATCO's workload measurement is by definition a controller-dependent measure (Prandini et al., 2011), which, depending on the number of controllers, their experience, etc., may be biased towards that particular group. Thus, in a quest of an objective sector complexity measure, the measurement technique should be flexible to adapt to changes not only in sector design but also the assigned controllers. The measure should represent the 'average' controller based on the 'average' mental capacity. Measures such as SD, DD, IO and also TLS, are all considered controller-dependent measures that are constructed using either controller workload measurements (SD, DD, and TLS approaches) or controller control activity measurements (IO approach).

In Europe, the sector's capacity estimation is based on the maximum number of aircraft that are to be controlled, while still permitting an acceptable level of controller workload (Majumdar et al., 2004). However, such a definition requires an understanding not only on controller workload and its measurement, but also a quantification of an acceptable level of controller workload or the threshold value at full capacity. As the static maximum number of aircraft represents only an indication of congestion, **traffic dynamic behavior** presents a more comprehensive picture of the 'real' situation. Figure 2.5 showed how a simple situation of nine aircraft with different heading orientation would create different complexity level.

An objective measure should be capable of presenting the effect of current air traffic behavior and also capture the **future traffic situation**, up to a certain prediction horizon. Of all the sector complexity measures discussed above, only the SD method does not include any future conditions of air traffic.

A summary of the various sector complexity measures is shown in Table 2.2. It is concluded that the SD, although it is widely used and gives good correlations with subjective workload ratings, does not fulfill any of the criteria. The proposed constraint-based metric based on SSD and the LE complexity map appear to be the most promising solutions. However, the LE presents its contribution through a display by showing the complexity map, whereas the SSD metric in this thesis functions as a measure through scalar value gathered from the SSD properties. In the next section, the construction of the SSD will be discussed in more detail.

Table 2.2: Criteria of an objective complexity measure (Adapted from Prandini et al. (2011)).

Metric	Independent of sector layout	Independent of controller	Captures traffic dynamic	Capture future condition	Output	
Static Density	No	No	No	No	Scalar value	
Dynamic Density	No	No	Yes	Yes	Scalar value	
Input-Output approach	Yes	No	Yes	Yes	Map	
Lyapunov Exponent	Yes	Yes	Yes	Yes	Map	
Traffic Load Smoother	No	No	Yes	Yes	Map	
Solution Space Diagram	Yes	Yes	Yes	Yes	Diagram	Scalar value

2.6 A Complexity Measure through an Obstacle Representation Method

The obstacle representation using velocity vectors has a long history in maritime navigation and motion planning. An overview of Collision Cone (CC) or Velocity Obstacle (VO) history starting from the introduction of obstacle representation dated as early as 1890s and the evolution of VO use in motion planning are discussed more extensively in Appendix A. The same principle that was used in CC (Chakravarthy & Ghose, 1998) and VO (Fiorini & Shiller, 1993) was adapted in ATC collision detection and avoidance approach. The aircraft collision detection and avoidance problem differs from most motion planning of obstacle avoidance problems studied in the robotics literature. This is due to the fact that the obstacles in ATC are never stationary (Goss et al., 2004).

A number of studies were being carried out actively in the ATC domain, in order to have a better traffic monitoring and also conflict detection, resolution and prevention capability. Obstacle representation and resolution possibilities using velocity vectors have been applied to the aerospace domain in numerous conflict detection and resolution studies. Bilimoria (2000), Goss et al. (2004) and Carbone et al. (2006) were a few who exercise the principle of CC in their researches. In studies conducted by Bilimoria (2000), an analytical geometric optimization approach to the problem of collision between two aircraft was presented, where they minimize the velocity vector changes required for conflict resolution of pair-wise encounter maneuvers (Bilimoria, 2000). Goss et al. (2004) have extended the research to a 3D environment using a mixed geometric and collision cone approach (Goss et al., 2004). The application of collision cone approach were later extended in research by Carbone et al. (2006) on a pair-wise non-cooperative decision for real-time applications.

2.6.1 The Basics of Solution Space Diagram Construction

In the effort to objectively measure complexity, the constraint-based approach based on the SSD, that uses similar obstacle representation methods as described above, has been suggested as a measure of sector complexity and ATCO taskload.

The SSD is used as a basis for obtaining a better metric for complexity, capable of measuring the taskload of air traffic controllers. The SSD can be described as the available control area for the controlled aircraft in respect to other observed aircraft within the vicinity. The construction of the SSD is based on the projection of the FBZ of the observed aircraft where the key constraint is the 5 NM separation minimum between aircraft. The construction of the SSD from the relative space to absolute space is illustrated in four steps, in Figure 2.10.

Consider a controlled (AC_{con}) and an observed (AC_{obs}) aircraft, as depicted in Figure 2.10a. The circle around AC_{obs} is the minimum separation that must be maintained. When tangent lines are drawn from either side of the minimum separation circle to the controlled aircraft position, a gray area as shown in Figure 2.10b appears. This area is called the FBZ. The velocity of AC_{con} relative to AC_{obs} should not lie inside the FBZ: if the relative velocity (V_{rel}) lays inside the gray area, both aircraft will experience separation violation in the future.

From the relative space, the FBZ can be projected to the absolute space by transposing the observed aircraft velocity vector to the SSD that consist of a minimum (V_{min}) and maximum (V_{max}) velocity possible for the controlled aircraft (aircraft performance limit) as seen in Figure 2.10c. It can be seen that the relative velocity related to the translated beam and the actual velocity of the controlled aircraft describe the same point in the diagram. Figure 2.10d presents the resulting SSD.

The SSD therefore presents the constraints on controlling AC_{con}, in terms of giving it heading and speed commands, caused by the presence of AC_{obs}. In Figure 2.10d, AC_{con} will lose separation with AC_{obs} when nothing is done. A turn to the left or right, or slowing down AC_{con} will resolve the conflict.

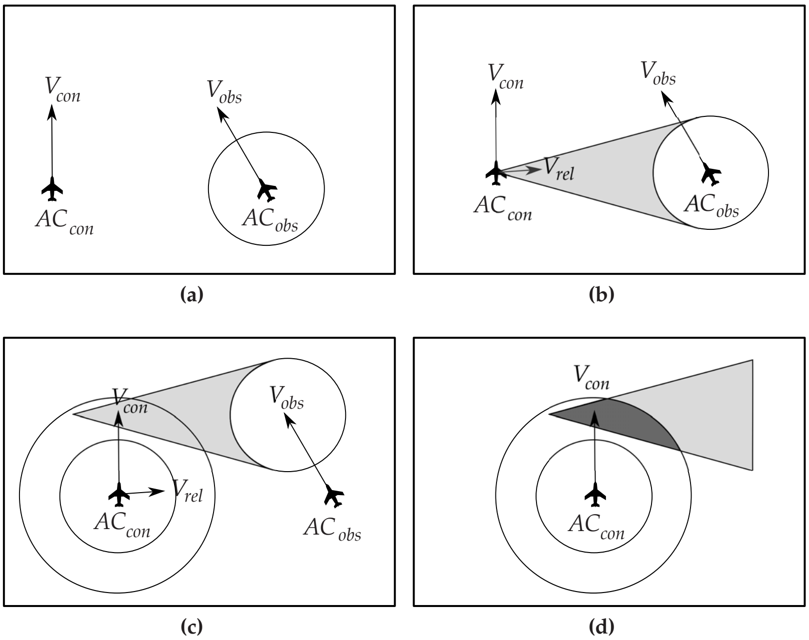


Figure 2.10: Procedure to construct the SSD. (a) Representation of observed and controlled aircraft with constant heading and velocity. (b) The corresponding FBZ of the observed aircraft. The area within the FBZ represents an instantaneous, complete set of possible conflicting relative velocities. (c) Translation of the FBZ on the SSD by transposing the observed aircraft velocity vector in order to made visible the relation between the relative and absolute space. The velocity vector is outlined based on an assigned look-ahead time. (d) Final SSD of single observed aircraft situation.

2.6.2 The Analytical Solution Space Diagram Construction

This thesis will use the analytical Solution Space Diagram (SSD) developed by Mercado Velasco et al. (2010) in exploring and measuring the sector complexity levels. There are three major steps in constructing the analytical SSD, which are: defining the collision trajectory, projecting the protected airspace circle to the collision trajectory and lastly, the envelope equation. The analysis conducted by Mercado Velasco et al. (2010) indicates that we could define the ''*Solution Space as a family of tangent circles with decreasing radius and centers on the collision velocity vectors*''. Figure 2.11 visualizes the steps followed to construct the FBZ.

The collision trajectory is the centerline of the FBZ, which defines a direct collision trajectory between the controlled and the observed aircraft. Therefore, this centerline

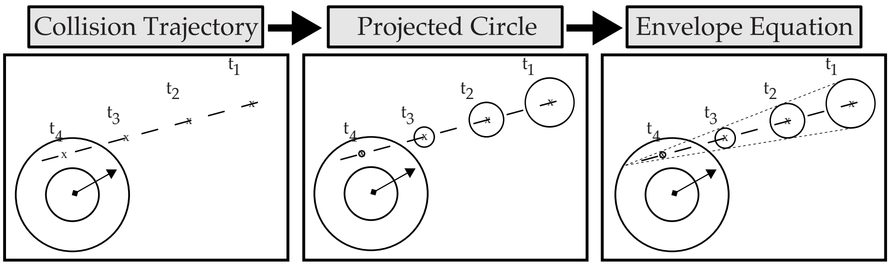


Figure 2.11: The envelope approach.

in the Solution Space diagram will represent velocity vectors that produce direct collisions.

The second step involves the making of the infinite series of projected circles centered on the centerline of the FBZ, which is the collision trajectory line and all circles sharing the same tangent lines. The projected circle originates from the protected airspace of every aircraft (5NM radius) with its center on the aircraft, where the radius of the projected circle is inversely proportional with look-ahead conflict time.

Lastly, the envelope equation of the family of circles is constructed, using the radius of the projected circle (projected protected airspace) of an observed aircraft for different conflict times.

2.7 Conclusion

This chapter discussed the foundations behind taskload, workload and the sector complexity measure. With workload and sector complexity being an intricate concept, it is difficult to evaluate both concepts by simply looking at the number of aircraft within a sector. While having the advancement of previous researches that also integrate the dynamic behavior of aircraft within the sector complexity measure and workload predictor, a measure which is independent of sector layout and controller differences is still at large.

In this thesis, an approach based on the investigations of problems using the SSD is suggested. Before elaborating further on more specific scenarios, an exploratory research will be done in the next chapter to investigate whether it is possible to recognize and predict high sector complexity levels. This will be elaborately discussed in order to fully understand the concept before applying it in the remainder of the thesis as a solution.

3

Solution Space as Sector Complexity Measure

In this chapter, the Solution Space Diagram (SSD), which shows all possible conflict-free vectors for aircraft, is proposed as a tool for assessing sector complexity. When considering the SSD for any individual aircraft, all neighboring aircraft introduce a zone of conflict, the Forbidden Beam Zone (FBZ) on the SSD. The changes in these FBZs are systematically studied in this chapter to increase the understanding of the SSD usability in measuring workload and sector complexity. The following sector variables are investigated: the aircraft intercept angle, aircraft speed and aircraft horizontal proximity. Static traffic simulations of two-aircraft conditions are set up for a number of case studies with different sector variables. These are then compared quantitatively through computing the area of the SSD. In this study it is assumed that a denser Solution Space results in a higher rating for the complexity factor. The results show that in cases where other variables are fixed to certain values, larger intercept angle and horizontal proximity produce a less dense Solution Space. On the other hand, higher observed aircraft speeds result in the FBZ being shifted outwards on the SSD. Whether or not it will induce higher or lower SSD area depends on the proximity and also the intercept angle of the observe aircraft.

3.1 Introduction

In order to enable controller's workload prediction or measurements, investigations in the area of sector complexity constructs becomes a priority. In this chapter the solution space approach is adopted to analyze, in a systematic fashion how sector complexity variables may have an impact on airspace complexity.

Previous research on sector complexity showed that the aircraft intercept angle (Remington et al., 2000, Rantanen & Nunes, 2005, Nunes & Kirlik, 2005, Lee et al., 2009), speed (Chatterji & Sridhar, 2001, Kopardekar & Magyarits, 2002, Rantanen & Nunes, 2005) and horizontal proximity (Laudeman et al., 1998, Sridhar et al., 1998) are some of the variables that have an effect on sector complexity.

The goal of the present study is to systematically analyze the properties of the Solution Space Diagram (SSD) due to changes in the sector complexity construct. It is hypothesized that by using these properties, we can obtain a better prediction of the sector complexity related taskload compared to existing methods.

Several sector variables are investigated using the SSD Plotter simulator (refer to Appendix B) namely: the aircraft intercept angle, aircraft speed and aircraft horizontal proximity. Other effects of certain traffic situation, such as number of aircraft and aircraft heading orientation to the SSD behavior will also be investigated. Quantitative analyses were conducted on the SSD area properties for the mentioned sector variables.

In the study of quantitative measurement of sector complexity, it is assumed that a denser Solution Space results in a higher level of sector complexity. This exploratory research is initially done in order to investigate whether it is possible to recognize (and with that, predict) high sector complexity levels. Experiments in a later stage will then verify the hypotheses gathered from this quantitative analysis and will provide a better understanding on the relationship between SSD area properties and workload as indicated by subjects and also related to controller's performance.

3.1.1 Static Traffic Simulation

To better understand the characteristics of the SSD, the SSD plotter was developed. The plotter enables the user to study the effect of any traffic situation on the SSD. The plotter consists of the Plan View Display (PVD) on the left side of the screen, three SSD of different aircraft on the right side of the screen and a number of buttons to navigate the controlled aircraft in the middle of the screen. Although more aircraft can be investigated using the SSD Plotter, the initial window setting with three aircraft present can be seen in Figure 3.1. Any speed and heading control can be done through the top SSD by selecting the desired heading or speed for the controlled aircraft. The functionality of the SSD Plotter will be further discussed in Appendix B.

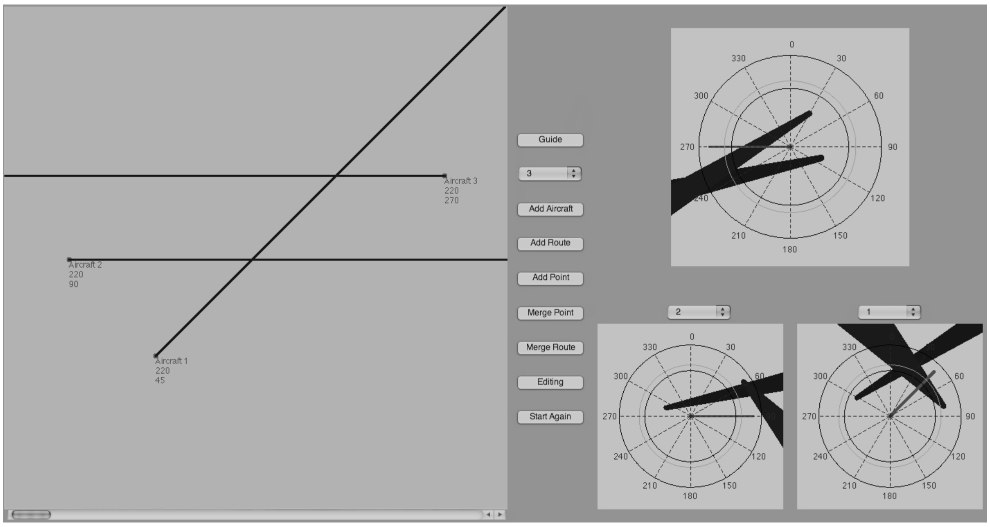


Figure 3.1: SSD Plotter.

3.2 Sector Complexity Variable Measure

The goal of this chapter is to evaluate several sector complexity construct effects on the SSD. The results gathered here are based on offline simulations of more than 100 case studies with various situations as detailed in the subsequent sections. To relate the taskload with non-solution space area, a continuous calculation of SSD area properties are being carried out.

In this case study, there are two area calculations method used, which are the whole unsafe area for aircraft i (A_{whole_i}) and the mean unsafe area (A_{mean}). Both area calculations are used in order to understand the effects of sector complexities on the available solution space.

The A_{whole} is calculated using the total area covered within minimum and maximum velocity-heading band (aircraft performance limit) of each individual aircraft. The A_{mean} is gathered using the sum of A_{whole} for all individual aircraft in the sector divided by the total number of aircraft. This will give an overview of the complexity metric for the whole sector.

$$A_{mean} = \frac{1}{n}\sum_{i=1}^{n} A_{whole_i} \tag{3.1}$$

While the A_{whole} might be an important SSD area property, which represents the constraints that limit each individual aircraft, the A_{mean} is a metric that represents the overall sector condition and will be used throughout the thesis. However, in the study conducted in this chapter, the A_{whole} will be used as a measure to analyze static two-aircraft conditions and the A_{mean} will be used as a measure to demonstrate the effect of different aircraft heading orientation for the same number of aircraft and position in a sector.

In a first attempt, we studied the effects of aircraft streams (that is, the airways or routes) intercept angles, the speed changes of aircraft on these streams and the horizontal proximity between aircraft on the SSD. For this purpose, several cases were studied. The cases that were being investigated always involved two intercepting aircraft (AC_1 and AC_2) at variable intercept angles, route lengths, and speed vectors. Quantitative analysis was then conducted on the SSD area properties for the mentioned sector variables.

Figure 3.2 shows an example of one of the case studies. Two converging aircraft are considered to merge to an intercept point. The sector variables studied are defined as follows:

1. **Horizontal proximity**: The Euclidean distance between aircraft, in NM.
2. **Intercept angle**: Angle of intercept on the intercept point of two aircraft, in degrees.
3. **Aircraft speed**: The velocities of both aircraft V_1 and V_2, in knots.

As sector complexity is inter-related, changes in any of these variables would result in changes in other related variables. Attention should be made towards changes that might occur in aircraft route length and angle between aircraft during the construction of different case studies. The aircraft route length is measured based on the aircraft initial position to the intercept point, in nautical mile (NM), while the bearing angle between aircraft is defined in degree. It is important to ensure that only one sector complexity variable is changed at a time in order to investigate the effects of a single variable change on the SSD.

The situation we elaborated here is based on two important assumptions. First, it is assumed that both aircraft have the same weight class and will have the same minimum and maximum velocities. Secondly, the minimum separation distance,

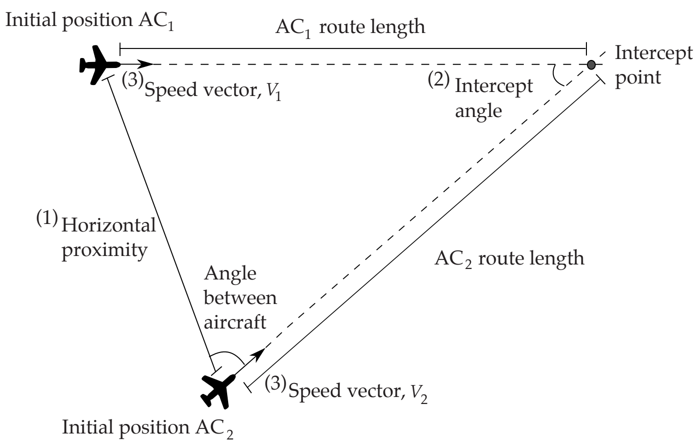


Figure 3.2: Example of case study properties.

represented by a Protected Zone (PZ) with radius of 5 NM around each aircraft, is to be maintained at all time. Using these settings, different sector complexity constructs are compared using a quantitative analysis in the next section.

3.3 Case Studies

The subsequent subsections will discuss a number of sector complexity constructs that may affect airspace complexity metric. Each section will discuss the behavior of SSD area properties in respect to changes in the sector variable.

3.3.1 Horizontal Proximity

Previous research on sector complexity has indicated that aircraft horizontal proximity is one of the variables that is responsible for the sector complexity construct. According to the Dynamic Density (DD) metric, aircraft that fly closer to one another have a larger weighting coefficient than those who are more distant (Laudeman et al., 1998, Sridhar et al., 1998).

In order to investigate the effect of horizontal proximities on the SSD, more than 50 position conditions with intercept angles of either 45°, 90° or 135°, were investigated. It is important to ensure that only one property is changed at a time. Figure 3.3 illustrates an example of two situations of different horizontal proximities. The velocity of both aircraft was maintained at the same value at all times. In order to maintain the angle between the two aircraft and to simulate situation of direct

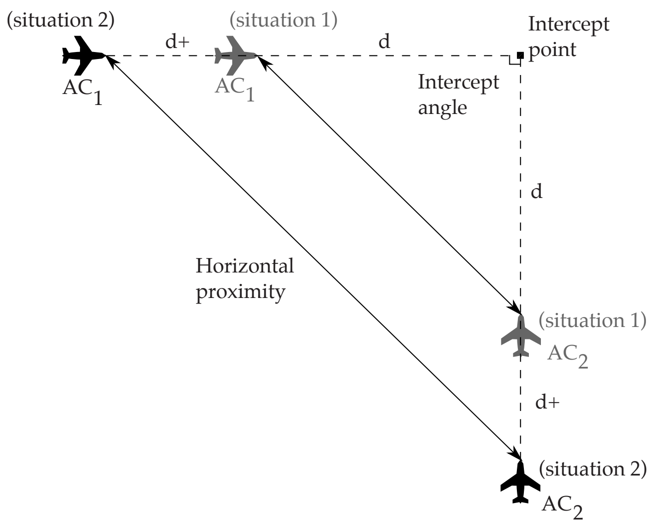
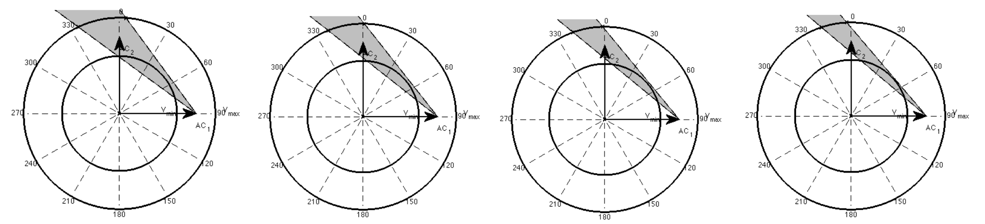


Figure 3.3: Horizontal proximity changes between aircraft with an intercept angle of 90°.

collision path, both aircraft have the same route length in both situations (d and $d+$ for situation 1 and situation 2, respectively). The effect of the horizontal proximity (at four instances) on the SSD is shown in Figure 3.4.



(a) (b) (c) (d)

Figure 3.4: SSD for AC_2 observing AC_1 at different horizontal proximities.

From the analysis, it was found that aircraft that are further apart from each other have a narrower Forbidden Beam Zone (FBZ) width than the ones being closer to each other. This can be seen in Figure 3.4 with aircraft progressing from being nearest (Figure 3.4a) to furthest (Figure 3.4d) apart from one another. The same pattern also applies to other intercept angles studied. The area covered is less dense for aircraft with a larger horizontal proximity where the area covered within the SSD decreases from 11% for the case in Figure 3.4a to 6% for case in Figure 3.4d. This also shows that a large horizontal separation between aircraft results in a less dense SSD, thus a lower complexity metric. A narrower width also implicates that there are more options to solve a conflict. This can be seen in Figure 3.4, where in Figure 3.4a and 3.4b, there is no option for AC_2 to resolve the conflict using a speed-only correction, whereas in Figure 3.4c and 3.4d the conflict can be resolved by either increasing or decreasing the AC_2 speed.

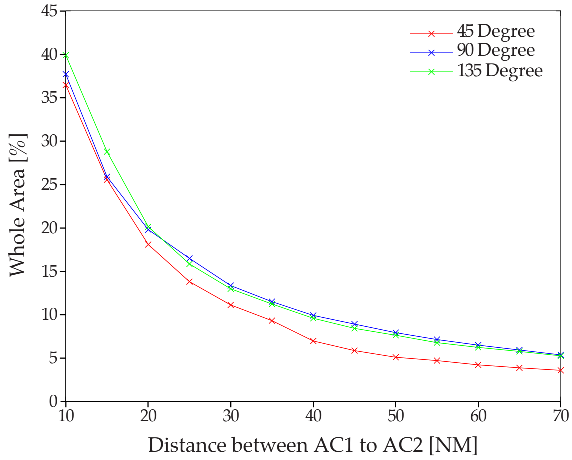


Figure 3.5: Effect of horizontal proximity based on different intercept angles.

Since both aircraft have the same route length with constants velocity, AC_1 will observe the situation in the same manner of AC_2. Thus, the same SSD area property is also observed for AC_1. Similar patterns were observed with different speed boundaries in conjunction with different intercept angles. Figure 3.5 illustrates the percentage area covered as a function of the horizontal distance and the intercept angle while maintaining the same velocity vector. It can be seen from this figure that the area properties decreases with larger distances between both aircraft at any intercept angle.

3.3.2 Intercept Angle

The ability of a controller to ascertain whether an aircraft pair would lose separation (more commonly known as conflict detection) is affected by a variety of variables that include, but are not limited to, the convergence angle (Remington et al., 2000, Rantanen & Nunes, 2005, Nunes & Kirlik, 2005). In order to understand the intercept angle as part of the sector complexity measure, the effect of intercept angle on the SSD area property is investigated.

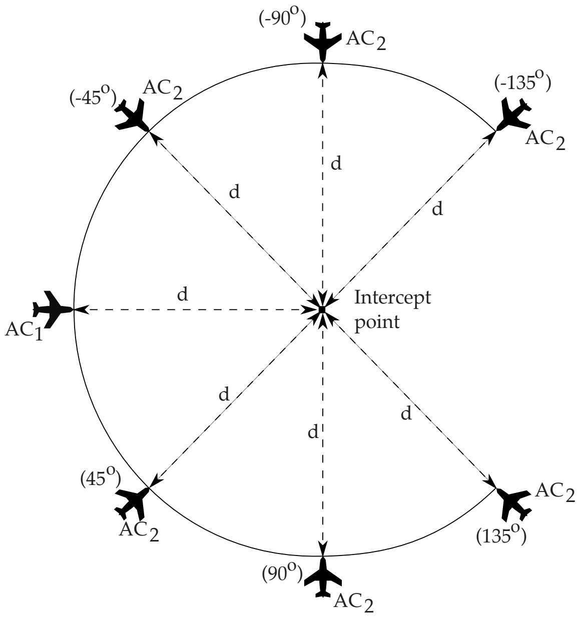
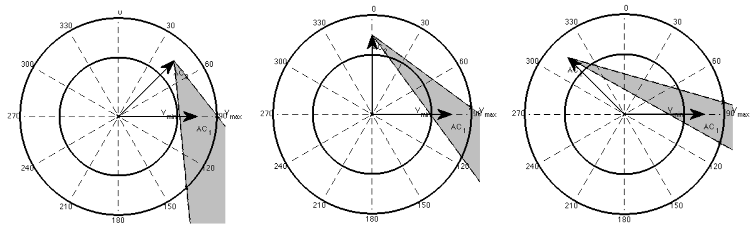


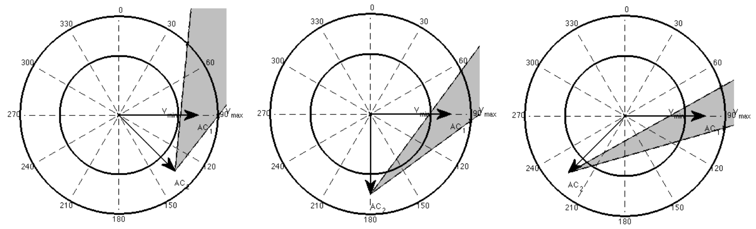
Figure 3.6: AC_1 observing intercept angle changes of AC_2.

The main goal is to investigate the effect of different intersection angles on the SSD. In the initial investigation, the route length between AC_1 and AC_2, d, remains constant and are equal at all times. Both aircraft were flying the same speed vector of 200 knots, but with different intercept angles between AC_1 and AC_2, which are 45°, 90°, 135°, −45°, −90° and−135°. The negative intercept angles were assigned for aircraft coming from the left, while positive intercept angles were assigned for aircraft coming from the right. As seen here, only changes in the intercept angles were investigated, while other variables were fixed to a certain value. Figure 3.6 illustrates the varying AC_2 positions, in order to produce different intercept angles between aircraft.

From the analysis, it is found that the larger the intercept angles of intersecting aircraft, the less dense the area within the SSD. Figure 3.7 shows the resulting SSD for different intercept angles. Figure 3.7 also shows the effect of aircraft coming from right (Figure 3.7a to 3.7c) or from the left (Figure 3.7d to 3.7e) side of the controlled aircraft. It is concluded here that aircraft coming from any direction with the same intercept angle and route length will demonstrate the same complexity measure due to the symmetrical nature of the conflict. For aircraft with 45°, 90° and 135° intercept angles, the SSD area properties are 14%, 11% and 8%, respectively. The same area properties hold for the opposite angle. This also shows that a larger intercept angle results in a lower complexity metric based on the properties of the SSD, because the non-solution space area covered with the conflict zone is smaller.



(a) AC_2 at 45° **(b)** AC_2 at 90° **(c)** AC_2 at 135°



(d) AC_2 at −45° **(e)** AC_2 at −90° **(f)** AC_2 at −135°

Figure 3.7: SSD for AC_1 observing AC_2 at different intercept angles.

However, this condition only applies if the observed aircraft is approaching from a certain direction, with equal or larger route length (indicated as d in Figure 3.6) than the controlled aircraft. When observing aircraft with d smaller than the controlled aircraft, a different pattern of SSD area was gathered. The effects of observing either 'front side' or 'backside' crossings then need to be elaborated.

Front side and Backside Crossings

It has been highlighted that there are differences between observing an aircraft crossing in front or from the backside of the controlled aircraft with an increasing intercept angle. Figure 3.8 illustrates several situations where AC_1 is observing front side and backside crossings at intercept angles of 45° and 135°. Both aircraft had the same speed of 220 knots and intercepted at the same point of the route.

In a case where the controlled aircraft (AC_1), located farther away from the intercept point, was observing an intercept of an aircraft (AC_2) crossing in front at a certain angle, the area covered was increasing with an increasing intercept angle. The area covered measured in a case of AC_1 separated at a distance of 30NM from the crossing point, observing AC_2 coming from a distance of 20NM was 3% for 45° intercept angle (Figure 3.9a) compared to 5% area covered for the 135° intercept angle (Figure 3.9b).

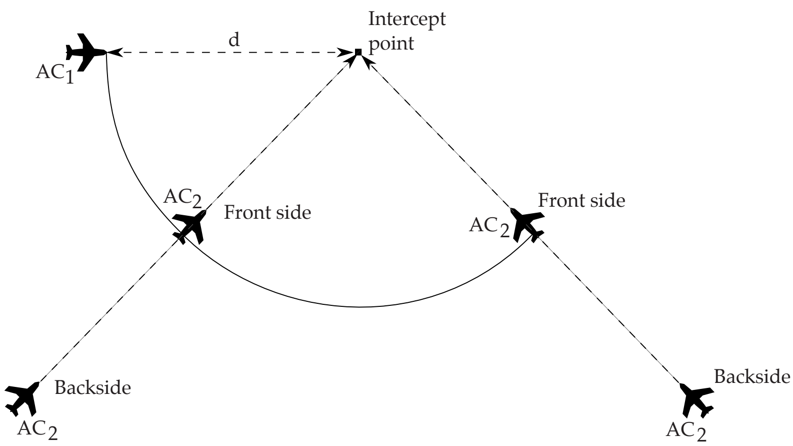


Figure 3.8: AC_1 observing front side or backside crossings.

When AC_1 is observing AC_2 crossing from the backside, the area covered was decreasing with increasing intercept angle. In this case, AC_1 is again at 30NM separation from the intercept point, observing AC_2 coming from a distance of 60NM from the intercept point. The area covered measured in this case is 8% for 45° intercept angle (Figure 3.9c) compared to 3% for 135° intercept angle (Figure 3.9d). These values indicate that a slightly higher complexity metric was found with an increasing intercept angle when AC_2 was already present in the sector and crossing AC_1 from the front side. The opposite situation occurs when AC_2 was approaching a sector and crossed AC_1 from the backside. It was concluded that an increase in intercept angle gave a lower complexity metric.

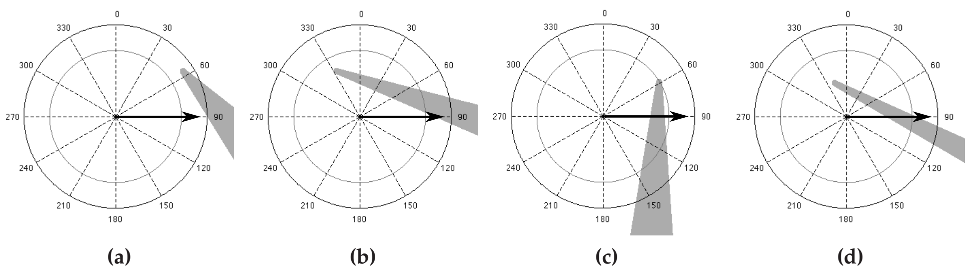


Figure 3.9: SSD for AC_1 observing AC_2 at different cases: (a) AC_2 crossing from the front side at 45°. (b) AC_2 crossing from the front side at 135°. (c) AC_2 crossing from the backside at 45°. (d) AC_2 crossing from the backside at 135°.

To extensively study the effect of intercept angle and the relative aircraft distance on the SSD area properties, several other cases were looked into and the results

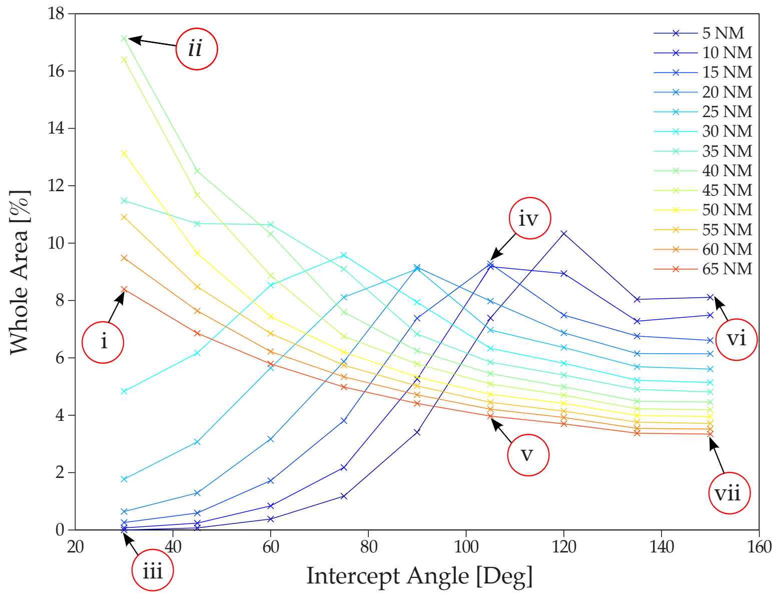


Figure 3.10: Plot of SSD area for AC_1 observing AC_2 with varying distances and intercept angles to the intercept point.

are illustrated in Figure 3.10. Several combinations of intercept angle and route length were numbered and the respective situations can be observed from Figure 3.11. Based on the initial study, it can be seen that observing present aircraft in the sector (with a distance from the intercept point less than 35 NM) will lead to an increase of SSD area with an increasing intercept angle. It is also observed that a larger intercept angle of incoming aircraft (aircraft with distance more than 35 NM) results in a less dense area inside the SSD with an increasing intercept angle. The results obtained here, matches the initial observations discussed earlier based on cases illustrated in Figure 3.8.

It is observed from Figure 3.10 that the SSD has an increasing area pattern from 5NM to 40NM route length, followed by a decreasing SSD area percentage. The same was observed for the 105° intercept angle cases. However, for the 150° intercept angle cases, it is observed that the SSD has a decreasing area from 5NM to 65NM route length.

Figure 3.11 shows an aircraft at 35 NM distance (d) from the intercept point, observing incoming or a present aircraft in the sector at a number of instances with different intercept angles and route lengths (with respect to the numbered points in Figure 3.10). The respective SSD plots can be observed from Figure 3.12 and 3.13.

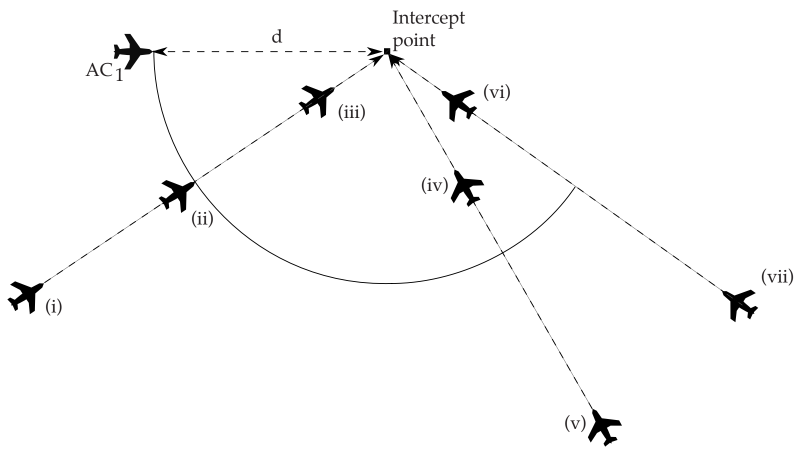
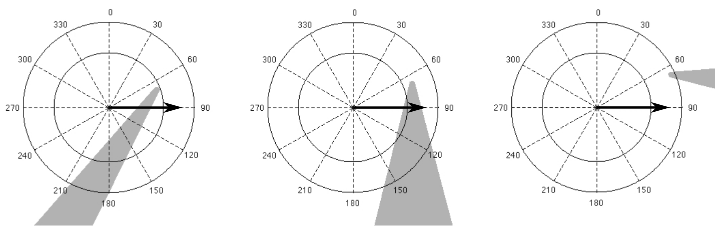
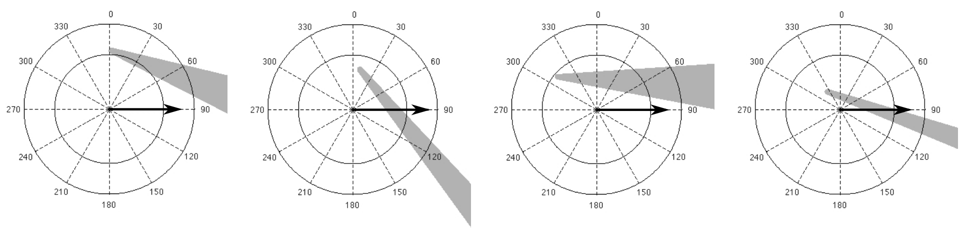


Figure 3.11: Position of AC_2 relative to AC_1 in varying cases of intercept angle and distance to intercept point.



(a) i **(b)** ii **(c)** iii

Figure 3.12: SSD for AC_1 observing AC_2 at three different cases with 30° intercept angle.



(a) iv **(b)** v **(c)** vi **(d)** vii

Figure 3.13: SSD of AC_1 observing AC_2 at four different cases with either 105° ((a) and (b)) or 150° ((c) and (d)) intercept angles.

The overview of SSD area properties with the effect of different intercept angle on the distance is illustrated in Figure 3.14. From the figure, it can be observed that a larger distance for larger intercept angles (120°, 135° and 150°) results in a continuing decrease of SSD area properties, thus relating to a lower complexity metric, whereas a larger distance for smaller intercept angles (30° to 105°) result in an initial increase of SSD area properties, thus relating to a larger complexity metric and followed by decreasing SSD area properties after a certain distance. This also suggests that for a larger intercept angle, the distance always relates to a less complex situation whereas for a smaller intercept angle, the increase of distance up to a certain point of route length relates to a more complex situation.

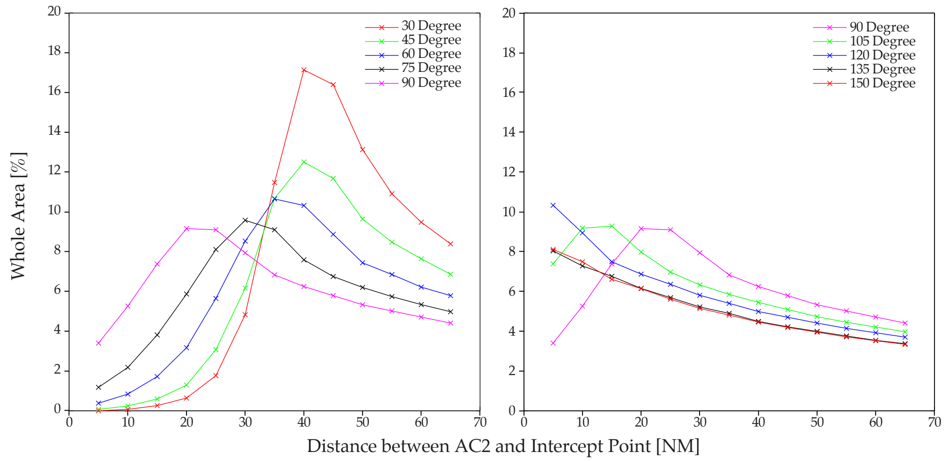
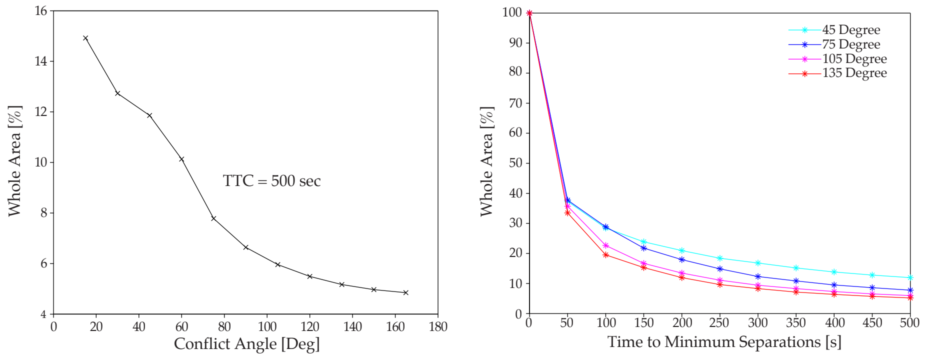


Figure 3.14: Plots of SSD area of AC_1 observing AC_2 at different distances and intercept angles relative to the intercept point.

Time to Conflict and Intercept Angle

The effect of intercept angle on the sector complexity construct was also investigated from a different perspective, namely the Time To Conflict (TTC). As illustrated in Figure 3.15a, with a fixed TTC at 500 seconds, a larger intercept angle will result in lower SSD area properties, thus a lower sector complexity construct. However, this is due to the larger distance between the aircraft for larger intercept angles, even with the same TTC value (refer to Figure 3.16). For the example in Figure 3.16, the intercept point is not on the track of AC_1 due to the application of 5NM separation minimum between aircraft.



(a) **(b)**

Figure 3.15: Plots of SSD area for AC_1 observing AC_2 in two different situations. (a) SSD area of AC_1 observing AC_2 at different intercept angle with constants TTC. (b) SSD area of AC_1 observing AC_2 at different TTC with constants intercept angles.

An example of the progression of a future conflict that will occur at an equal time in the future with different intercept angles is shown in Figure 3.15b. Based on Figure 3.15b, a larger intercept angle results in lower SSD area properties.

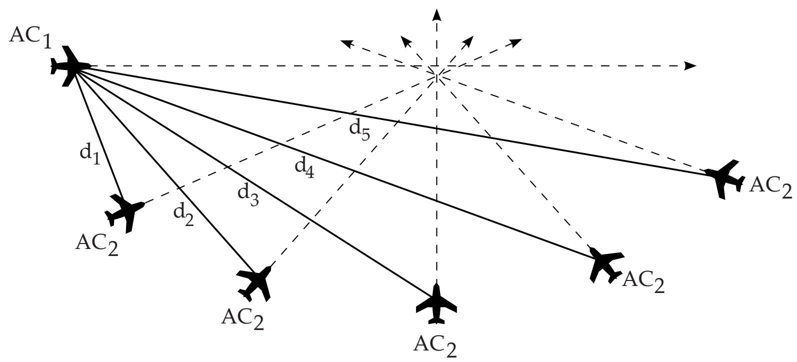


Figure 3.16: AC_1 observing AC_2 at different intercept angle with the same TTC.

3.3.3 Aircraft Speed

A previous study by Rantanen & Nunes (2005) suggested that speed is a confounding factor to conflict or intercept angles and the ability to detect a conflict. It has an impact on whether the estimation must be done on the distance (aircraft traveling at same speed) or the time (aircraft traveling at different speeds) to the point of trajectory intersection Rantanen & Nunes (2005). Research by Kimball et al. (1973) reported that increasing the speed difference between converging objects results in a lower accuracy. This is due to the fact that the controller now has to integrate two (rather than one) pieces of speed information and project their implications. This shows the importance of studying the effect of speed variations to the sector complexity, especially when coupled with the intercept angle.

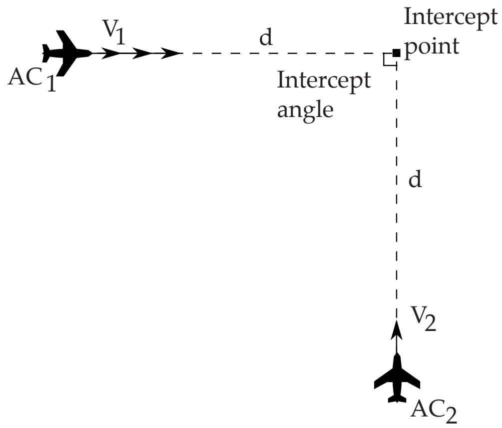
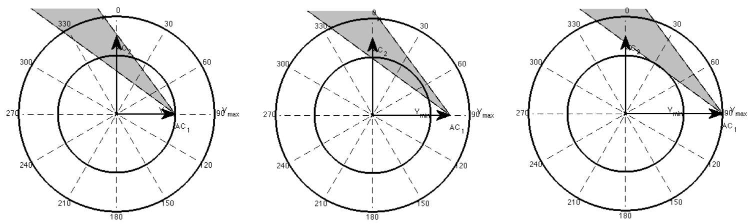


Figure 3.17: AC_2 observing speed changes of AC_1 at 90° intercept angle.

A number of cases of aircraft pairs at the same distance between each other were investigated. Figure 3.17 illustrates a case where change of speed only occur for AC_1. The resulting SSD can be observed in Figure 3.18, where the speed and heading of the observed aircraft can be seen on the SSD of the controlled aircraft through the position of the tip of the FBZ. This is because the FBZ is obtained by transposing the triangular conflict zone with the observed aircraft velocity vector.



(a) V_1 = 150 knots **(b)** V_1 = 200 knots **(c)** V_1 = 250 knots

Figure 3.18: SSD for AC_2 observing AC_1 at three different speeds of AC_1 with constant aircraft position.

In Figure 3.18, AC_1 will encounter a separation violation problem in the future with AC_2 when the aircraft maintains its current heading and speed. However, giving speed or heading instructions to one or both aircraft can resolve the future separation issue. In this case, an increase (Figure 3.18a) or decrease (Figure 3.18c) in speed for AC_2 will solve the future separation issue. It might not be desired for on-course aircraft to change the heading angle in order to fulfill efficiency constraints, however, if it is required to maintain safety, it may be the proper way to resolve a conflict, such as in Figure 3.18b. It is found that the higher the speed of the observed aircraft, the more the FBZ in the SSD is shifted outwards. The changes in the speed only affect the currently controlled aircraft's SSD. Because there is no change of speed for the controlled aircraft, AC_2, the corresponding diagram for AC_1 observing AC_2 remains the same during the change of speed vector in AC_1.

The whole area covered on the SSD depends on the relative positions and the intercept angle of both aircraft, where a shift outwards of the FBZ will be translated as more or less SSD area percentage covered. This can be seen by comparing Figure 3.18a to 3.18c where a shift outwards results in more area covered within the SSD, which gives the value of 8%, 11% and 15% area covered for cases in Figure 3.18a, 3.18b, and 3.18c, respectively. Hence it can be hypothesized that larger relative speeds can result in a higher or lower complexity metric, depending on the position and intercept angle of the aircraft.

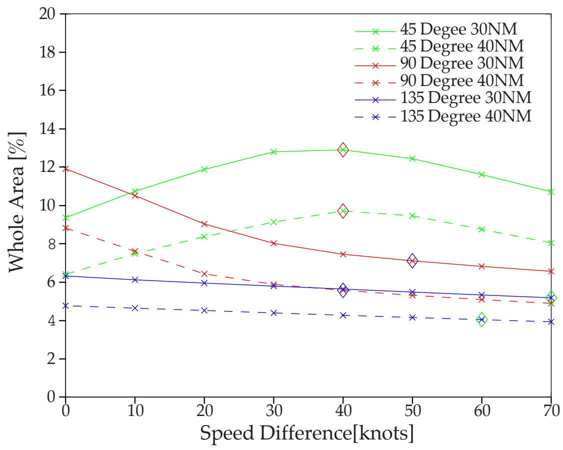


Figure 3.19: SSD area for AC_2 observing AC_1 speed change with varying intercept angle (aircraft position is constant).

The effect of speed changes was also investigated further for aircraft intercepting at 45°, 90° and 135° with more possible cases, and the results are illustrated in Figure 3.19. Differences in intercept angle, speed limit band (which may represent differences in aircraft performance limits or aircraft types) and the size of the speed limit were investigated. Figure 3.19 shows the effect of speed changes on a 180 - 250 knots speed band, with both AC_1 and AC_2 at either 30 NM or 40 NM distance from the intercept point at different intercept angles. Both aircraft initial speeds were 250 knots, and to illustrate the effect of speed variations, one of the aircraft was given a gradual speed reduction toward 180 knots.

The diamond shapes in Figure 3.19 indicate the minimum difference needed for aircraft to avoid future separation violation. Based on Figure 3.19, the effect of speed and distance is evident with 45°, 90° and 135° intercept angles showing a decrease in the SSD area properties with a larger relative distance while maintaining the trends of the graph. In 90° and 135° cases, larger distances also indicated that a smaller speed difference (marked with diamond) was needed in order for both aircraft not to be in a future separation violation. Figure 3.19 also shows that aircraft flying at a smaller intercept angle need less speed difference than aircraft flying larger intercept angle to avoid future separation violation caused by having the same flight path length to the intercept point.

On the other hand, the effect of the intercept angle shows different patterns in SSD area properties in regards to the speed variations. It can be observed from Figure 3.19 that a 45° intercept angle showed an increase of SSD area properties up until the intermediate speed limit followed by a decrease of SSD area properties with decreasing aircraft speed. However, for 90° and 135° intercept angle cases, the reduction of speed is followed by a continuing decrease in SSD area properties.

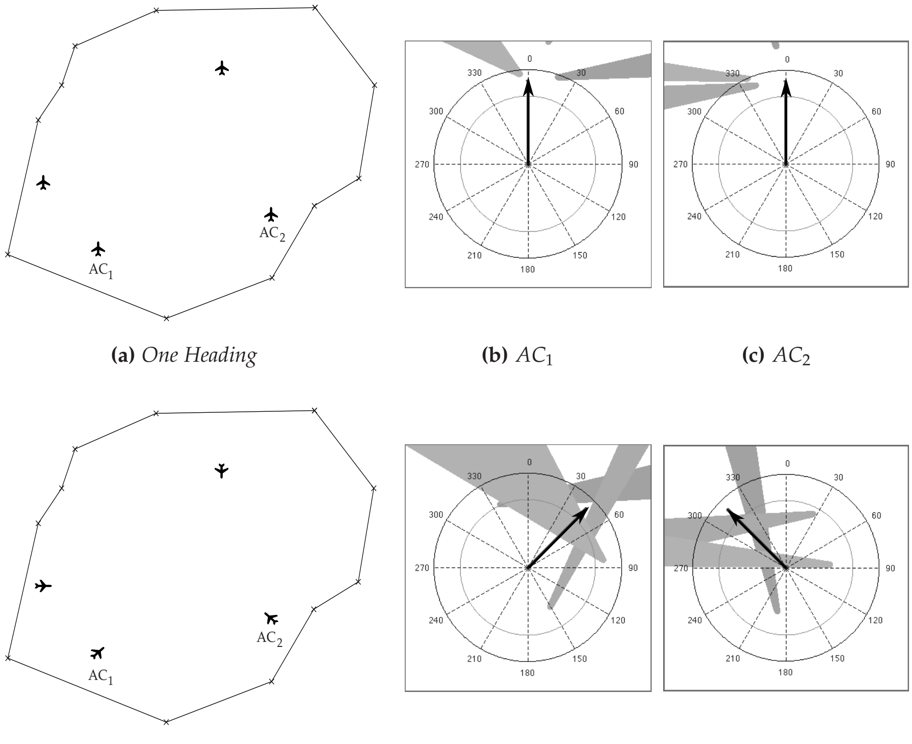
3.3.4 Number of Aircraft and Aircraft Heading Orientation

One of the most popular and easy-to-use methods to measure sector complexity is through the measurement of aircraft Static Density (SD). It is commonly used to obtain an instantaneous indication of the sector complexity. It is defined as, the number of aircraft per unit of sector volume. However, representing sector complexity solely based on the SD measure has its shortcomings by not being able to sufficiently capture the geometry and dynamics of situations. This section discusses this issue using a simple example of the effects of the number of aircraft and it's heading orientations, and how it is captured by the SSD area properties.

Figures 3.20 and 3.21 show examples of the number of aircraft and the aircraft heading orientation that was investigated here. The SSD for AC_1 and AC_2 as specified in Figures 3.20 and 3.21 were illustrated for all cases. All aircraft within the sector are free of conflicts.

In a four-aircraft situation, illustrated in Figures 3.20a and 3.20d, an A_{mean} of 9% and 16%, respectively, were gathered. Whereas in a six-aircraft situation, illustrated in Figures 3.21a and 3.21d, an A_{mean} of 15% and 20%, respectively were gathered.

Based on the SSD area properties, it was clear that more aircraft in the sector relates to a higher SSD area properties, comparing cases in Figure 3.20a to Figure 3.21a. The corresponding SSD also illustrates the effect of adding two aircraft to the SSD of AC_1 and AC_2 where two additional FBZ were present in Figure 3.21b and 3.21c if compared to Figure 3.20b and 3.20c.

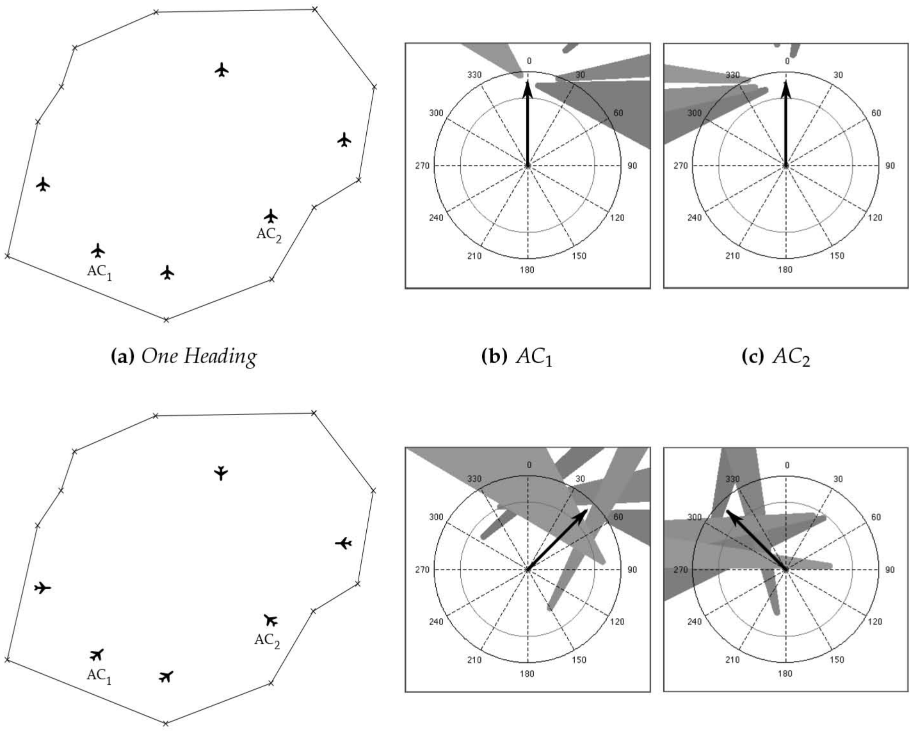


(a) *One Heading* **(b)** AC_1 **(c)** AC_2

(d) *Multiple Heading* **(e)** AC_1 **(f)** AC_2

Figure 3.20: Different heading orientations for four aircraft located in a sector with the same position.

This case study also agrees with the notion that aircraft heading orientation also influences the complexity construct of a sector through cases illustrated in Figure 3.20 and Figure 3.21. Here it can be seen that cases with converging aircraft (Figure 3.20d and Figure 3.21d) result in much higher SSD area properties than cases where all aircraft have an equal heading (Figure 3.20a and Figure 3.21a). The SSD also shows the effects of heading with Figure 3.20b and 3.20c showing the FBZ of aircraft with one heading and Figure 3.20d and 3.20e showing the FBZ of aircraft with several headings. The same four-aircraft situation in Figure 3.20 and six-aircraft situation in Figure 3.21 shows to have a higher complexity with several aircraft headings.



(a) *One Heading* **(b)** AC_1 **(c)** AC_2

(d) *Multiple Heading* **(e)** AC_1 **(f)** AC_2

Figure 3.21: Different heading orientations for six aircraft located in a sector with the same position.

The SSD area properties of the situation in Figure 3.20d (A_{mean} of 16%) and Figure 3.21a (A_{mean} of 15%) also demonstrates that more aircraft does not necessarily mean higher complexity, but that the heading orientations of aircraft within the sector matters more.

3.4 Conclusion

The SSD represents a constraint-based method, which uses the areas covered by the FBZ conflict areas as an indication of the level of difficulty of the situation that a controller has to handle. When the FBZs on the SSD occupy more area, fewer possible solutions are available to resolve future separation violations. This chapter aimed at deliberating the effects of various sector complexity construct properties on the SSD area properties. In the analysis, it is assumed that a denser area is related to a higher complexity.

The sector complexity of varying horizontal proximity between aircraft, intercept angle, aircraft speed, as well as the number of aircraft and aircraft heading orientation were illustrated through the SSD area percentages. It is important to ensure that only one property is changed at a time.

From the initial study conducted, it is concluded that higher intercept angles, result in a smaller complexity metric, in a case where the observed aircraft has a route length larger than or equal to the controlled aircraft. For horizontal proximity properties, it was found that aircraft, which are further apart, have a lower complexity metric. This is true for all intercept angles. A larger aircraft speed, on the other hand, may result in higher or lower complexity metric, with respect to the observed aircraft position and intercept angle.

To investigate the effect of the number of aircraft within a sector (aircraft Static Density (SD)), some general scenarios were investigated. Based on the result, having more aircraft within a sector results in a higher SSD area percentage. However, it is also highlighted that the sector complexity is not only affected by the number of aircraft present in the sector, but also how the aircraft's headings are orientated.

It should be noted that because of the dynamic behavior of the airspace's complexity, these sector complexity parameters did not change individually at each instant. As an introductory chapter of SSD behavior towards sector complexity construct variables, this case study will provide the basis for hypotheses that will be tested systematically in subsequent studies. To further understand the behavior of the SSD, it is important to investigate other and more combinations of sector complexity metrics. In the following chapters, the findings regarding the relationship between sector complexity factors and SSD metrics will be validated by means of human-in-the-loop experiments to also obtain the Air Traffic Controller's insight on the perceived workload and how this can be related to the SSD area properties.

4

Solution Space in Merging Scenarios

This chapter studies the effect of several sector complexity variables on workload in dynamic merging scenarios. A number of sector properties were investigated, namely number of streams to be merged, the merge angle, the proximity of incoming aircraft and the variability of the traffic mix of small and large aircraft. The effect of these sector design variables on Air Traffic Controller (ATCO) workload and also Solution Space Diagram (SSD) were evaluated and compared in a human-in-the-loop experiment with two groups of subjects. Based on the findings, each sector complexity variable led to a different effect on both workload and SSD area properties. However, conclusive relation between workload and SSD behavior cannot be drawn as the findings gathered in the experiment only reveals either only workload or SSD area properties to be significant for single sector complexity factor. It is also observed that a change in one sector complexity variable would result in other sector complexity variable to also change. This has highlighted the issue of designing single sector complexity factor in a dynamic environment. Having said that, the overall correlation between the workload and the SSD area properties was found to be higher than correlation to the number of aircraft within the sector, showing that the SSD could be a good workload predictor.

4.1 Introduction

In the effort to better understand the effect of sector complexity variable on workload, an experiment was conducted in which traffic patterns were varied during a merging scenario. In a human-in-the-loop experiment, the Solution Space Diagram (SSD) was used as an offline evaluation method of sector complexity and workload, and metrics based on the SSD were compared to workload ratings given in the experiment.

Led by findings gathered in the initial theoretical study (Chapter 3) combined with the results from several research projects on what determines sector complexity (Laudeman et al., 1998, Sridhar et al., 1998, Remington et al., 2000, Rantanen & Nunes, 2005, Nunes & Kirlik, 2005, Lee et al., 2009), the experiment looked at several variables that are responsible for sector complexity. Based on an initial quantitative study conducted, it is concluded that different sector variables produce different sector complexity patterns (Chapter 3).

This experiment is aimed at investigating whether the behavior gathered in static situations will also be present in dynamic situation. We focused on the effect that sector complexity has on workload, and verify whether the SSD is a good predictor for workload. Four sector variable were explored in this experiment, namely (1) the number of streams to be merged, (2) the merge angle, (3) the proximity of incoming aircraft and (4) the variability of the traffic mix of small and large aircraft. To do that, each controller was presented with six scenarios of 20 minutes. Subjects were assigned to merge aircraft with a predetermined route, and have aircraft leave the sector at a speed of 180 knots. Two different groups of controllers participated in this experiment: a group with experience (expert) and a novice group (student). It was expected that the differences between expert and non-expert subjects' behavior should be detected.

The result gathered on the effect of the aforementioned sector complexity variables towards workload and the SSD area properties will be detailed. Then, together with the SSD, a number of other performance related measures, such as the number of aircraft delivered, Root Mean Square (RMS) Extra Distance Ratio and RMS Exit Speed were evaluated in order to investigate the level of correlation between both SSD area properties and subject's performance with the workload rating in varying scenario properties situation.

4.2 Experimental Design

4.2.1 Subjects and Tasks

Twelve male subjects participated in the experiment. There were two groups of test subjects, representing two population groups: six Graduate Students of the

Faculty of Aerospace Engineering and six subjects that participated in an extensive Air Traffic Control (ATC) introductory course. These two population groups are named in this research as Students and Experts, respectively. The Experts ages ranged between 26 and 48 (average age, μ = 35.80, standard deviation, σ = 9.50), and the Students ages ranged between 24 and 33 (μ = 27.40, σ = 3.51). Every subject had a scenario sequence assigned to him. The application of two groups of subjects was to investigate possible differences in control strategies between both groups and also to study the robustness of the method applied.

This experiment was set-up using a standalone simulator. Figure 4.1 shows the interface that was used in the experiment. The left part of the display shows the Plan View Display (PVD) area, which shows the route, the sector under control, the surrounding area of the sector and the aircraft within the area.

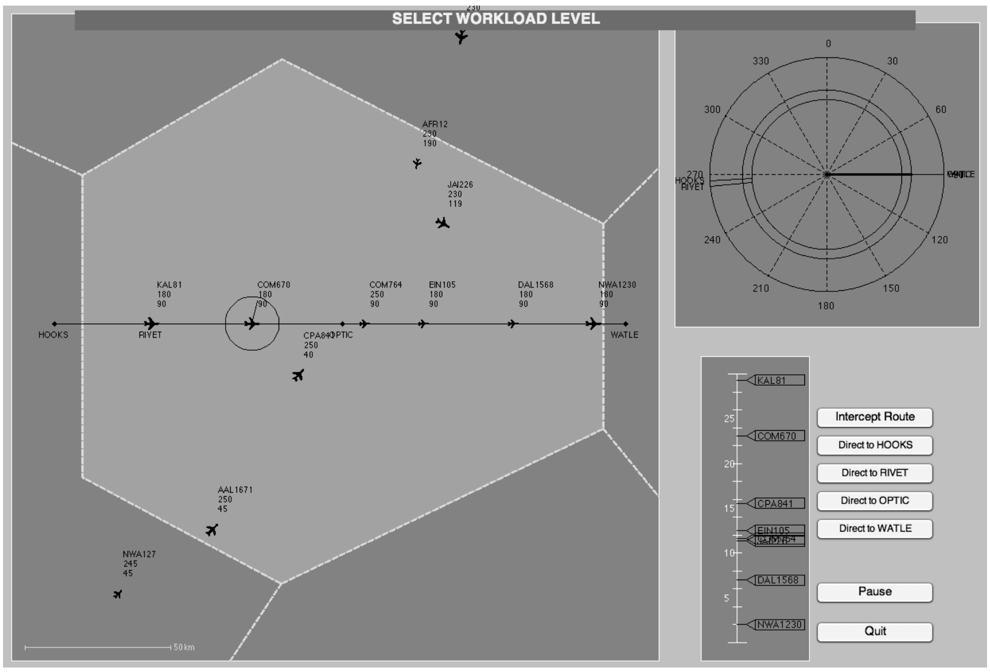


Figure 4.1: Experiment simulator.

The right part consists of the control panel area where heading and/or speed commands, 'Direct To' and 'Intercept Route' commands can be made to selected aircraft. The outer and inner circles on the heading and/or speed command area correspond to the maximum and minimum speed of the selected aircraft, respectively. The middle circle represents the 180 knots assigned exit speed. A thick line represents the current speed and heading vector of the aircraft. During the experiment, controller can instruct an aircraft to change its speed and/or heading using mouse control.

The 'Intercept Route' button was used to instruct an aircraft to maintain current heading and speed until the route is crossed and then to continue along the route. The 'direct to' buttons were used to direct an aircraft to fly to a particular route point and from then on to follow on the route.

Information regarding the remaining time it would take an aircraft that was instructed to fly on the route to reach the exit point, i.e., WATLE, was shown in the bottom right of the simulator in Figure 4.1. The time information is only visible when the aircraft is assigned to the route or the sectors exit point and it is labeled with the aircraft callsign.

The subjects were asked to merge all incoming aircraft to the predetermined route through intercepting with the route or by directing the aircraft to one of the four route points, namely HOOKS, RIVET, OPTIC and WATLE. During the experiment, the participants were asked to rate their perceived workload every 60 seconds. An automated stimulus provided a message on the display (Figure 4.1) that triggered the participants to rate their workload by means of typing a number between 1 (low workload) and 7 (high workload) on the keyboard.

4.2.2 Independent Variables

The independent variables in the experiment are (1) the number of streams, which is either 3 or 4 streams, (2) the intercept angle, which is either 45° or 100°, (3) the traffic mixes of heavy and light aircraft and (4) the proximity between incoming aircraft streams, which also has two levels (high and low proximity). In the high proximity scenarios, both incoming streams comes from north west and north east of the sector (Figure 4.2a), whereas in low proximity scenarios, the two incoming streams comes from north east and south west of the sector (Figure 4.2b). The independent variables provide a total number of 6 experiment conditions that is shown in Table 4.1.

Table 4.1: Scenario settings.

Scenario	No. of Streams	Streams Setting	Proximity	Angle	% of Light Aircraft
I	3	Main route, +45°, +100°	High	-	43%
II	3	Main route, −45°, +100°	Low	-	39%
III	4	Main route, +45°, −45°, +100°	-	45°	40%
IV	4	Main route, +45°, +100°, −100°	-	100°	43%
V	3	Main route, +45°, +100°	High	-	33%
VI	3	Main route, −45°, +100°	Low	-	32%

where the + / - angle is defined from the main route

The presentation order of the scenarios were randomized, in order to counterbalance a possible order effect in the experiment. A total of six scenarios (indicated in the table as I to VI) with six incoming aircraft sequence (indicated in the table as S1 to S6) and six traffic mix settings (indicated in the table as T1 to T6) were presented to each subjects. The combinations of independent variables in the full scenarios are shown in Table 4.2. The following sections will further elaborate the sector layout and settings.

Table 4.2: Sequence of scenario based on subject.

Subject	Group	Sequence											
		1		2		3		4		5		6	
1	EXPERT	I	S1T3	II	S2T2	V	S5T6	III	S3T1	VI	S5T6	IV	S4T4
2	EXPERT	II	S2T2	III	S4T4	I	S6T3	IV	S3T1	VI	S1T5	V	S5T6
3	STUDENT	III	S3T4	IV	S4T1	II	S6T2	V	S5T5	I	S2T3	VI	S1T6
4	STUDENT	IV	S3T1	V	S5T6	III	S4T3	VI	S1T5	II	S6T4	I	S2T2
5	EXPERT	V	S5T5	VI	S2T6	IV	S4T4	I	S6T3	III	S3T2	II	S1T1
6	STUDENT	VI	S2T5	I	S1T3	V	S5T6	II	S6T1	IV	S3T4	III	S4T2
7	EXPERT	IV	S4T1	V	S1T5	I	S6T2	VI	S2T6	III	S3T3	II	S5T4
8	STUDENT	V	S1T6	II	S6T3	IV	S2T5	III	S3T4	IV	S4T1	I	S5T2
9	STUDENT	I	S2T1	IV	S3T3	III	S4T2	V	S1T6	II	S5T4	VI	S6T5
10	STUDENT	V	S6T6	VI	S5T5	IV	S3T2	II	S1T4	I	S2T1	III	S4T3
11	EXPERT	II	S6T4	III	S4T2	VI	S5T5	I	S1T3	V	S2T6	IV	S3T1
12	EXPERT	III	S3T2	I	S5T1	II	S2T4	IV	S4T3	VI	S6T6	V	S1T5

4.2.3 Dependent Measures

During the experimental runs, there are a number of dependent measures that were being monitored in order to get more insight into the subjects' performance in relation to the workload rating given.

i) Subject's Performance: Three subject's performance measures were measured in the experiment. Firstly, the number of aircraft delivered that represent the total number of aircraft delivered per subject per scenario. Secondly, to determine whether subjects complied with the instruction to ensure that aircraft leave the sector at 180 knots, the Root Mean Square (RMS) Exit Speed was measured. It calculates the value of the differences between the exit speed of all delivered flights in every scenario to the assigned exit speed of 180 knots. Lastly, the RMS Extra Distance Ratio, which calculate the difference between the most efficient trajectory (that would be a straight line from the point at which the aircraft enters the sector to WATLE) to the trajectory generate by the subject. The value of the extra distance ratio of all delivered flights per subject and scenario are calculated as follow:

$$\eta_{ijk} = \frac{d_{ijk} - d_{ijk_{min}}}{d_{ijk_{min}}}$$

where d_{ijk} is the actual flown distance of flight k, controlled by subject i in scenario j, and is the minimum possible distance for that flight.

ii) Subject's control activity: Three control activity were measured during the experiment. Firstly, the number of commands ($N_{Command}$) counted as the total number of speed and/or heading command given by subject. Secondly, the number of 'direct to's' given by subject. Lastly, the number of airspace click (N_{Air}) counted as the total number of clicks on the airspace section. Despite the fact that a click on the control panel area or a direct to's can have the effect of giving two commands (heading and speed changes), we assume that one click cannot represent a cognitive effort worth of two commands.

iii) SSD area properties: The SSD area properties was calculated using the mean of SSD unsafe area (indicated in this chapter as SSD) of all aircraft within the sector ('$SectorArea$') or based on the whole projected area ('$CompleteArea$'). It is gathered using the following equation with A_{whole} representing the total area within minimum and maximum velocity-heading band of each individual aircraft and n being the number of aircraft. The SSD area properties were measured every 60 seconds to match with the workload rating instances.

$$A_{mean} = \frac{1}{n} \sum_{i=1}^{n} A_{whole_i}$$

iv) Number of aircraft properties: Two type of aircraft count were measured by the total number of aircraft within the currently controlled sector (N_{Sect}) and total number of aircraft which are present on the screen presented to the subject. (N_{All}).

v) Subject's difficulty ratings: The workload rating, measured on a one to seven scale, were provided by the subject every 60 seconds during the experiment run. In order to correct for inter-subject differences, Z-scores of the subjective ratings were used in the subsequent data exploration. This correction was performed by calculating the Z-scores for every test subject.

4.2.4 Sector Layout

A total of six scenarios were set up in the experiment (referred in Table 4.2 as I to VI). Each scenario had distinct characteristics in order to represent different sector complexity variables. In an attempt to gather more information of the dynamics of events, to avoid under-achievement from the controllers, and following the procedure of previous experiments (Hermes et al., 2009, d'Engelbronner et al., 2010, Mercado Velasco et al., 2010) the simulation was run as a fast-time simulation at 4 time of real-time. In previous experiments, where this technique was applied, no negative effects were reported. With this simulation speed, every scenario was run for 20 minutes real-time, with some breaks in between scenarios.

Incoming Aircraft

There were four incoming aircraft patterns tested in the experiment: two patterns with three streams and two with four streams. The incoming aircraft patterns for the three and four streams scenarios are presented in Figure 4.2.

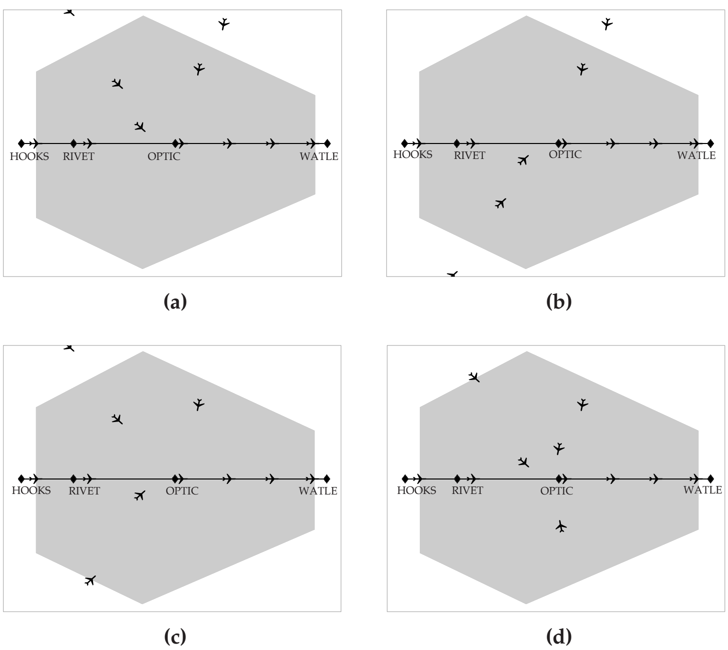


Figure 4.2: Incoming aircraft pattern. (a) Three streams (Scenario I and V) (b) Three streams (Scenario II and VI) (c) Four streams (Scenario III) (d) Four streams (Scenario IV).

Traffic Mix and Conditions

The traffic density level in every scenario was established with the arrival of new aircraft at fixed intervals of time. Training scenarios had one aircraft every 240 seconds simulated time, while full scenarios had one aircraft every 190 seconds simulated time. In the initial setup of the experiment, there will be eight aircraft during the training session and 11 during full session.

Six traffic mix situations (indicated in Table 4.2 as T1 to T6) were defined to get the required percentage of light and heavy aircraft mixture between 30% and 43% of light aircraft during each scenario. Two conditions were defined having light aircraft ranges between 32% and 33% of the total aircraft (Scenario V and VI) and four conditions were defined having light aircraft between 39% and 43% of the total aircraft (Scenario I to IV).

The motion of aircraft was simulated using simplified kinematic equations. A constant acceleration (or deceleration) of 3 m/s^2 (5.83 $knots/s$) and a constant heading change rate of 3 deg/s were used to compute trajectories either for heavy or light aircraft. The same will be used in the experiments in Chapter 5 and 6.

Scenario Combination

By combining the six scenarios, four sector complexity variables could be investigated as shown in Table 4.3. Two cases of proximity and number of streams were investigated. In proximity case, one with higher light aircraft traffic mix (between 39% and 43%), and one with lower light aircraft traffic mix (between 32% and 33%) and in number of streams case, one with almost the same percentage of light aircraft (between 39% and 40%) and one with mixed percentage of light aircraft (between 32% and 40%). This is to see the effect of different combination of proximity and number of streams towards traffic mixes. More combinations could be found, however, only the respected six cases shown in Table 4.3 were investigated.

Table 4.3: Scenario setting combinations.

Case	Scenario Combination	
	A	B
Proximity (Case 1)	Scenario I: High Proximity	Scenario II: Low Proximity
Proximity (Case 2)	Scenario V: High Proximity	Scenario VI: Low Proximity
No. of Streams (Case 1)	Scenario II: 3 Streams	Scenario III: 4 Streams
No. of Streams (Case 2)	Scenario VI: 3 Streams	Scenario III: 4 Streams
Traffic Mix	Scenario I: 42% light Aircraft	Scenario V: 32% light Aircraft
Intercept Angle	Scenario III: 45°	Scenario IV: 100°

4.2.5 Procedure

Subjects were briefed on the nature of the experiment, the goals to be achieved and the simulator that was used for the experiment. Following that, one training session was conducted in order to get the test subject familiarized with the simulator. After the test subject indicated that they understood the task to be performed, a total of six full experiment scenarios were presented to the test subjects according to the scenario sequence planned. Only the six experiment sessions were considered to be the measurement sessions of which the recorded data will be used in the analysis. A single run of a condition lasts approximately 20 minutes.

4.2.6 Hypotheses

The experiment was intended to study the effect of sector complexity variable during a merging scenario to the controller's workload. It was also the intention of the experiment to study and confirm the relation between SSD area properties and workload. It is hypothesized that the SSD area properties are related to Air Traffic Controller (ATCO) workload. It is also hypothesized that a certain configuration of these complexity variables leads to different effects on workload and also SSD area properties. Some of the hypotheses are based on the findings gathered in the Chapter 3 of the effects of traffic situations on the SSD. The hypotheses concerning complexity metric are as follow:

1. **Proximity**: Higher proximity between incoming aircraft streams result in a more dense area of the SSD (from Chapter 3) and therefore higher complexity metrics or workload.
2. **Number of Streams**: Having a larger number of incoming aircraft streams impose higher taskload on subjects.
3. **Traffic Mixes**: Having a mix of traffic (light and heavy aircraft) implies that different speed limits are present. Having a larger traffic mix would impose higher taskload on subjects.
4. **Intercept angle**: Larger intercept angle of observed aircraft crossing from the backside of the controlled aircraft results in a less dense area of the SSD and therefore lower complexity metrics or workload. However, larger intercept angle of observed aircraft crossing in front of the controlled aircraft results in a more dense area of the SSD (from Chapter 3) and therefore higher complexity metrics or workload.
5. **SSD as a predictor of workload:** It is hypothesized that the SSD area properties can predict controller's workload in a merging task with higher SSD area properties revealing a higher controller workload.

4.3 Results

In this section, the effect of several sector complexity variables changes towards workload and SSD area properties is discussed. The exploration of different sector complexity variables was carried out in order to study the effect it has on workload and the ability of SSD to capture it. In this result section, effects were considered significant at a probability level $p \leq 0.05$, where p is the probability that the null hypothesis is true.

4.3.1 Proximity between Incoming Aircraft Streams

Firstly, the effect of proximity to workload was investigated through different aspect, namely the workload rating, subject performance and also SSD area properties. A pattern of subject's behavior towards proximity of incoming aircraft streams were gathered based on the findings. Two different proximity situations were tested, high and low incoming aircraft proximity.

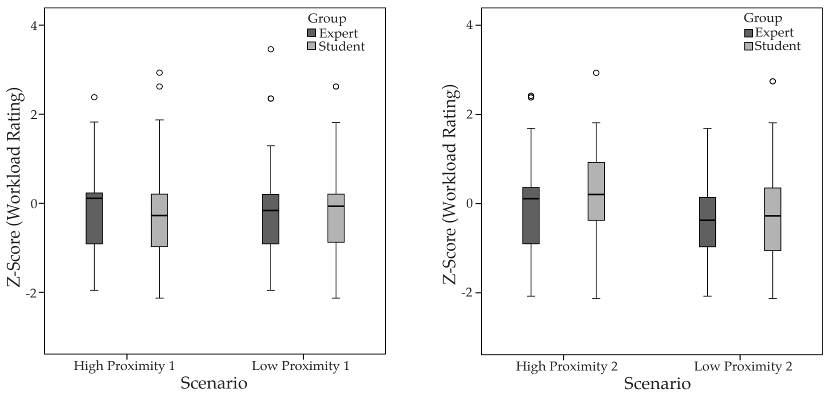
Workload Rating

The analysis was performed on the average z-scored workload rating from a 20 minutes experiment run. This was done to gather an overall workload situation of each individual scenario. Analysis on the significance of proximity of incoming aircraft using Wilcoxon-Signed Rank test (test statistic Z) showed no significant results (Table 4.4). The findings are also illustrated in Figure 4.3 that shows the pattern of workload ratings based on incoming aircraft proximity.

Table 4.4: Wilcoxon-Signed Rank test on workload rating based on incoming aircraft proximity.

	Case 1		Case 2	
	Student	Expert	Student	Expert
Z	-0.734	-0.314	-1.572	-1.782
p	0.463	0.753	0.116	0.075

In the case of proximity of incoming aircraft, based on Table 4.4, lower p-values were present for both groups in Case 2. Both groups also show the same median value trend with high incoming aircraft proximity cases, showing higher median workload ratings (Figure 4.3). This could suggest a behavior of a higher workload level for higher proximity cases between incoming aircraft. It is also interesting



(a) Case 1 **(b)** Case 2

Figure 4.3: Workload rating based on incoming aircraft proximity.

to note that in the same scenario routing, differences in traffic mixes can result in different workload ratings. It is assumed that difference in the percentage of light aircraft (43% and 39% in Case 1 compared to 32% and 33% in Case 2) might have influenced the behavior of subjects during the experiment and in the end resulted in a different scenario picture and taskload. However, the p-values are not significant enough to be conclusive. To fully confirm or reject the hypothesis, further exploratory experiment with larger sample sizes are needed.

Subject's Performance

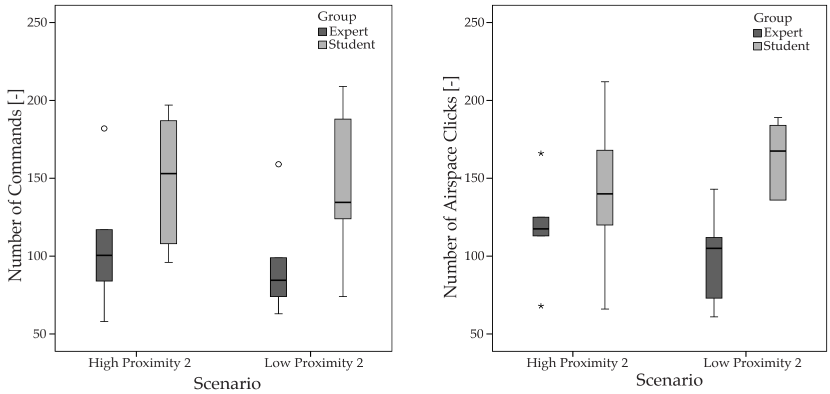
Based on the analysis, it is revealed that all three subject's performance metrics (the number of aircraft delivered, RMS Extra Distance Ratio and RMS Exit Speed), showed no significant effects between high or low proximity of incoming aircraft.

Subject's Control Activity

Another parameter that was looked into is the subject's control activity such as the $N_{Command}$ and N_{Air}. Based on the results in Table 4.5, the control activity between high and low incoming aircraft proximity in Case 2 was seen to be significantly different only in group expert. It is observed that in Case 2 the expert group have lower $N_{Command}$ and N_{Air} in low incoming aircraft proximity cases compared to high incoming aircraft proximity cases. This can be observed from Figure 4.4. Student group on the other hand are not influenced in terms of control activity for different cases of horizontal proximities and percentages of light aircraft.

Table 4.5: Wilcoxon-Signed Rank test on control activity based on incoming aircraft proximity.

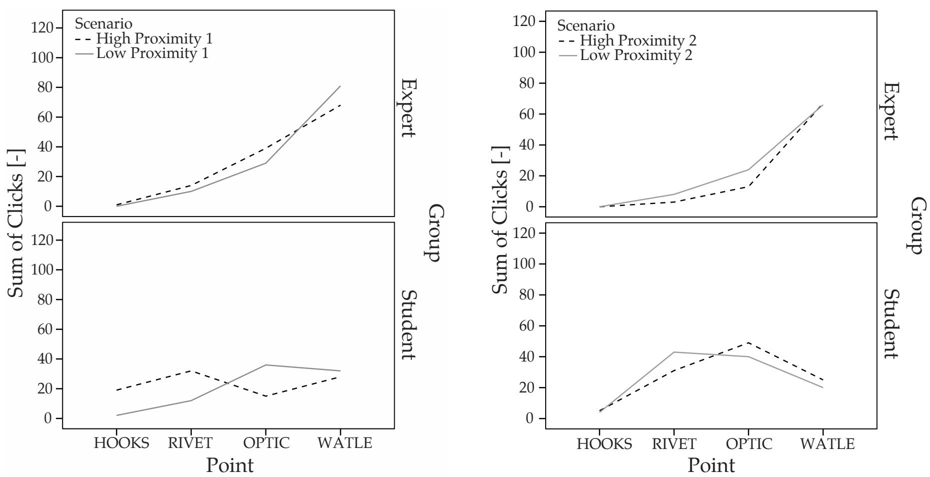
	$N_{Command}$				N_{Air}			
	Case 1		Case 2		Case 1		Case 2	
	Student	Expert	Student	Expert	Student	Expert	Student	Expert
Z	-0.210	-1.156	-0.631	-1.997	-0.631	-1.682	-0.734	-2.201
p	0.833	0.248	0.528	**0.046**	0.528	0.093	0.463	**0.028**



(a) $N_{Command}$ **(b)** N_{Air}

Figure 4.4: Control activity based on incoming aircraft proximity (Case 2).

However, when applying a Bonferroni correction with significance cutoff at 0.025 for post-hoc analysis, no significant effect were found in the control activity based on different proximity between incoming aircraft streams.



(a) Case 1 **(b)** Case 2

Figure 4.5: Direct to's activity based on incoming aircraft proximity.

One of the elements in the command behavior is the direct to's towards one of the available route point. Figure 4.5 elaborates on the pattern of direct to's used by the different groups of subjects. Overall, the way the groups behaved in giving direct to's command was different between one another. In both cases, the expert group showed an increasing number of command to the final route point, whereas for student group, all route points were treated almost the same in term of preference (Case 1) or with higher preference for the two middle route point (Case 2). This shows different styles of separation assurance between groups where experts uses both vector and speed to ensure separation with intercepting directly to the final route point and students mostly use speed commands to ensure separation with intercepting to the route and giving the aircraft proper spacing.

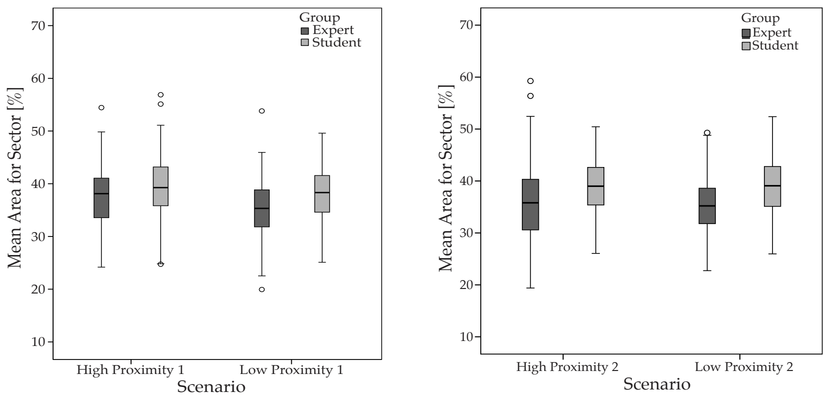
Solution Space Diagram Area Analysis

Similar to the workload rating analysis, here, the analysis was also performed on the average SSD area properties gathered from a 20 minutes experiment run. Based on the analysis, all of the area properties that are of significance showed the same trends in the mean rank values with a larger area found in high incoming aircraft proximity cases. The results of SSD for both Sector and Complete Area were shown in Table 4.6. The values that were statistically significant were highlighted.

Table 4.6: Wilcoxon-Signed Rank test on the SSD for Sector and Complete Area based on incoming aircraft proximity.

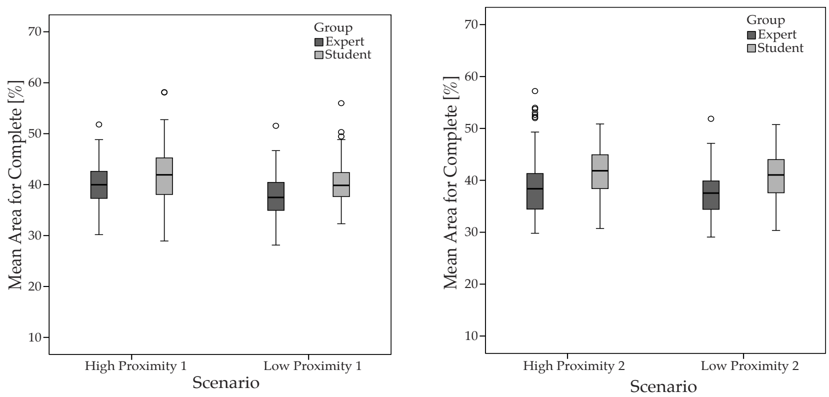
	Sector Area				Complete Area			
	Case 1		Case 2		Case 1		Case 2	
	Student	Expert	Student	Expert	Student	Expert	Student	Expert
Z	-0.734	-1.782	-1.992	-0.943	-1.572	-1.992	-1.153	-0.943
p	0.463	0.075	**0.046**	0.345	0.116	**0.046**	0.249	0.345

Based on the results, Sector Area presentation is shown to have a significant effect in Case 2 for students. On the other hand, Complete Area presentation is shown to have a significant effect in Case 1 for experts. Figure 4.6 and 4.7 show the SSD area properties for Sector and Complete Area, respectively. For both significantly affected cases, high incoming aircraft proximities result in higher SSD mean area properties and this can be seen from Figure 4.6 and 4.7. However, the same could not be confirmed after post-hoc analysis by applying a Bonferroni correction significance cutoff at 0.025.



(a) Case 1 **(b)** Case 2

Figure 4.6: SSD area properties based on incoming aircraft proximity (Sector Area).



(a) Case 1 **(b)** Case 2

Figure 4.7: SSD area properties based on incoming aircraft proximity (Complete Area).

4.3.2 Number of Streams

The effect of number of streams to workload, subject performance and also SSD area properties was also investigated. Two different sector number of streams with either three or four streams situation were investigated in this section. The same data treatment as used in previous section, where the average workload rating and SSD area properties, calculated from the 20 minutes experiment duration was maintained.

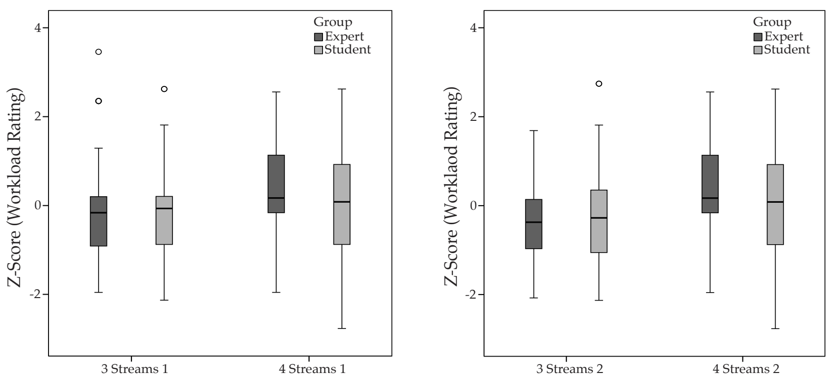
Workload Rating

Results of Wilcoxon-Signed Rank test on workload rating are shown in Table 4.7. Case 1 and Case 2 showed a significant effect on expert's workload rating with Z = -1.992, p = 0.046 (for both cases). The corresponding workload rating is illustrated in Figure 4.8 with significance value demonstrating larger difference in workload rating values. It can be observed from Figure 4.8 that more incoming aircraft streams results in a higher workload rating.

Table 4.7: Wilcoxon-Signed Rank test on workload rating based on number of streams.

	Case 1		Case 2	
	Student	Expert	Student	Expert
Z	-0.734	-1.992	-1.153	-1.992
p	0.463	**0.046**	0.249	**0.046**

Differences in the way both groups handle traffic, very much effect the dynamics of the scenario. Experts were observed to have different tactical solution for different number of streams. As a result, a significantly different workload rating were gathered based on different number of streams scenarios. This will be explained further in the analysis on the subject's control activity section.



(a) Case 1 **(b)** Case 2

Figure 4.8: Workload rating based on number of streams.

Figure 4.8 illustrates the findings in regards to workload rating. Based on the figures, both experts and students showed the same pattern of workload rating where higher workload rating in four streams cases were found compared to three streams cases. These findings, however was not confirmed with post-hoc analysis using a Bonferroni correction significance cutoff at 0.025.

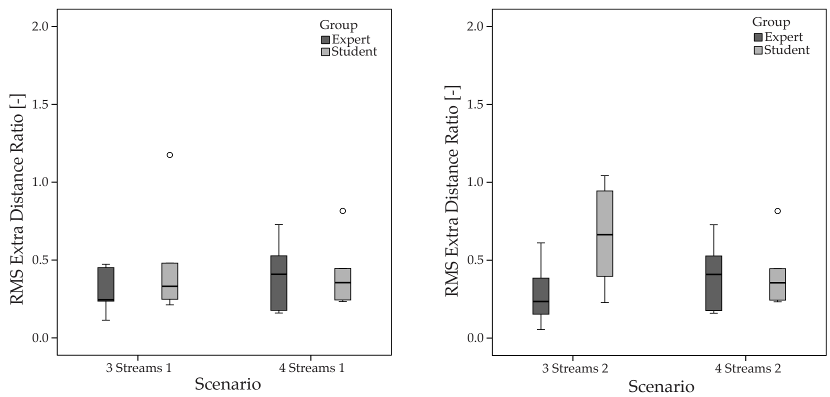
Subject's Performance

Subject's performance was analyzed in order to investigate the effect of number of streams. The only performance metric that showed significant effect based on difference in number of streams was RMS Extra Distance Ratio. The summary of performance metrics statistical analysis can be seen in Table 4.8.

Table 4.8: Wilcoxon-Signed Rank test on performances based on number of streams.

	RMS Extra Distance Ratio			
	Case 1		Case 2	
	Student	Expert	Student	Expert
Z	-0.524	-1.572	-1.992	-1.153
p	0.600	0.116	**0.046**	0.249

The RMS values were chosen instead of the average Extra Distance Ratio, to be able to capture possible magnitude of varying extra distance value. Thus, it is not the difference in average extra distance ratio per scenario that is of interest, but how much it varies over delivered aircraft. With this, the performance during the entire run can be observed.



(a) Case 1 **(b)** Case 2

Figure 4.9: RMS Extra Distance Ratio based on number of streams.

Based on the results, only the student group showed a significantly different RMS Extra Distance Ratio values. Closer inspection of the performance values in Figure

4.9 revealed that for student group, a higher RMS Extra Distance Ratio values were observed in three streams case compared to four streams case. This is evident only in Case 2. It is assumed that differences in percentages of light aircraft present in Case 2 might have affected the behavior of the student group. However, the same could not be confirmed after post-hoc analysis by applying a Bonferroni correction significance cutoff at 0.025.

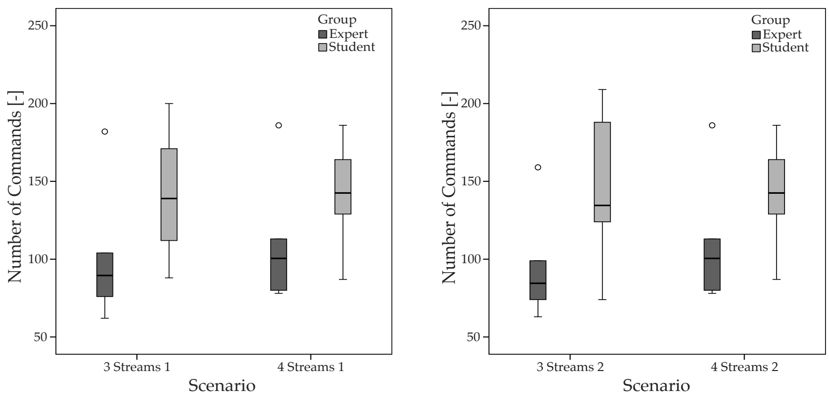
Subject's Control Activity

Analysis on the control activity pattern revealed that the expert subjects showed a significance effect of number of streams on control activity for both cases. Student group, on the other hand showed a significance effect only in Case 1. Summary of analysis results can be observed in Table 4.9. Figure 4.10 and 4.11 illustrates the significant behavior gathered from the analysis.

Table 4.9: Wilcoxon-Signed Rank test on control activity based on number of streams.

	$N_{Command}$				N_{Air}			
	Case 1		Case 2		Case 1		Case 2	
	Student	Expert	Student	Expert	Student	Expert	Student	Expert
Z	-0.105	-2.201	-1.992	-2.201	-1.051	-2.201	-2.201	-2.201
p	0.917	**0.028**	**0.046**	**0.028**	0.293	**0.028**	**0.028**	**0.028**

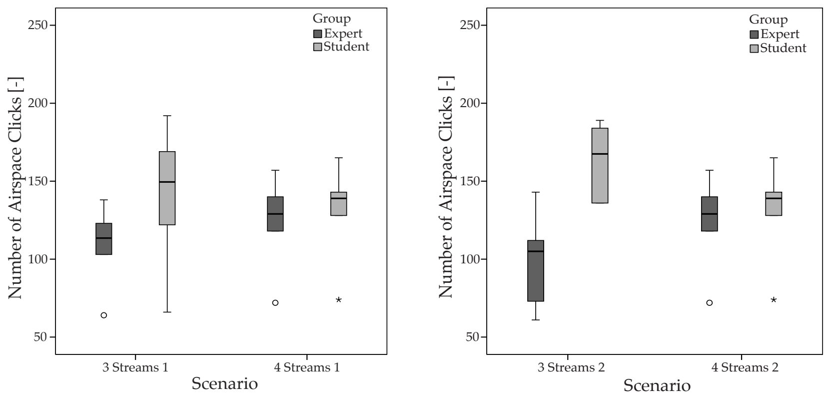
It should be made aware that in every scenario, the incoming aircraft was fixed to one aircraft every 190 seconds simulated time. This setting ensured that the number of aircraft within a sector remains comparable even with different number of streams.



(a) Case 1 **(b)** Case 2

Figure 4.10: $N_{Command}$ based on number of streams.

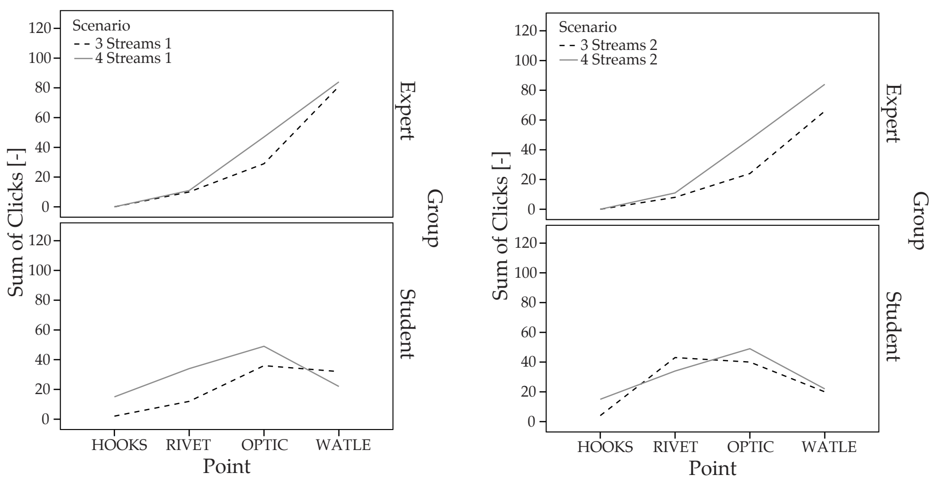
Based on Figure 4.10 and 4.11 it can be observed that higher $N_{Command}$ and N_{Air} can be seen in four streams cases compared to three streams cases. Although the pattern is only valid for expert where it is observed that for student, less N_{Air} were found in four streams cases compared to three streams cases. However, even when the activity of choosing aircraft to control is lesser in four streams cases, the total $N_{Command}$ is still higher.



(a) Case 1 **(b)** Case 2

Figure 4.11: N_{Air} based on number of streams.

However, post-hoc analysis by applying a Bonferroni correction significance cutoff at 0.025 revealed that the effect seen in control activity based on different number of stream in a sector is insignificant.



(a) Case 1 **(b)** Case 2

Figure 4.12: Direct to's activity based on number of streams.

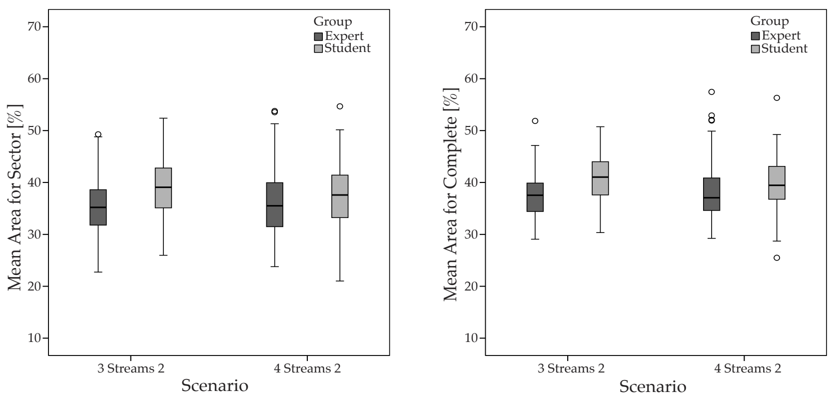
Another command-related activity, namely direct to's pattern was also analyzed in order to see the preference of route owned by each group. Based on the findings it is again observed that experts prefer to direct the aircraft to the final route point in contrast to student group which prefer the middle two route point more than the final route point (Figure 4.12). It is also observed that even when the expert group consistently prefer the final route for both scenarios, higher sum of $N_{Command}$ and N_{Air} is observed in four streams cases.

Solution Space Diagram Area Analysis

Analysis on the SSD area reveal that no significant effect was found for both cases of different number of streams. Results of the analysis is shown in Table 4.10. However, to discuss the behavior of the SSD area properties in varying number of streams, plots from Case 2 results, which has a lower p-values were illustrated.

Table 4.10: Wilcoxon-Signed Rank test on the SSD for Sector and Complete Area based on number of streams.

	Sector Area				Complete Area			
	Case 1		Case 2		Case 1		Case 2	
	Student	Expert	Student	Expert	Student	Expert	Student	Expert
Z	-0.734	-1.153	-1.153	-1.363	-0.524	-0.524	-1.153	-1.363
p	0.463	0.249	0.249	0.173	0.600	0.600	0.249	0.173



(a) Sector Area **(b)** Complete Area

Figure 4.13: SSD area properties based on number of streams (Case 2).

Figure 4.13 illustrate the behavior of SSD area properties with different number of streams for the two subject groups. Based on the figures, it is observed that fewer streams results in slightly higher SSD area properties. This is caused by the condition with four streams producing a more disperse traffic pattern than condition with three streams. An example of this situation is illustrated in Figure 4.2. In both situation, the incoming aircraft sequence was at the same rate with four streams having more option to alternate. However, the difference between the SSD area properties are not significant enough to be captured in the statistical analysis. Thus, the behavior of the SSD metrics' behavior cannot be compared to the workload rating's behavior.

4.3.3 Traffic Mixes

The effect of the traffic mix on workload was investigated by means of taskload rating, performance metrics and control activity and also SSD area properties. Two different traffic mixes with either 41% light aircraft or 32% light aircraft were investigated in this section. Again, the analysis in this section was performed by taking the average workload rating and SSD area properties calculated from the 20 minutes experiment duration.

Workload rating

Wilcoxon-Signed Ranked analysis showed that the traffic mixes did not results in any significant effect on the workload rating for both group of controllers with Z = -1.572, p = 0.116 and Z = -0.674, p = 0.500 for student and expert group, respectively. However, the box plot of controller workload rating based on different traffic mixes is illustrated in Figure 4.14 to observe differences in workload ratings.

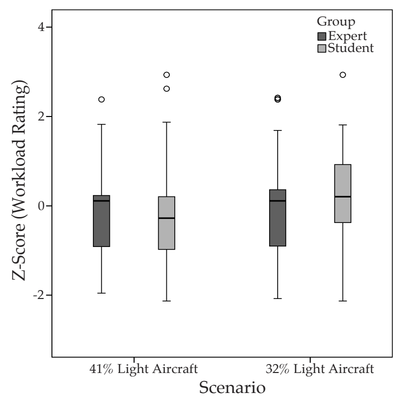


Figure 4.14: Workload rating based on traffic mixes.

Based on the plots, the traffic mix showed a larger difference in workload ratings in the student group. It was also indicated only by the students at the end of every scenario runs that they noticed the different types of aircraft more, because they used the minimum velocity to separate the aircraft in order to get a good merging sequence. Based on Figure 4.14, it shows that higher median values (workload ratings) are seen when the traffic has fewer light aircraft present, thus confirming the comments given by the subjects. However, these differences were not significant enough with the current sample size to neither confirm nor reject the hypothesis.

Subject's Performance

Based on the analysis, it is revealed that all three subject's performance metrics (the number of aircraft delivered, RMS Extra Distance Ratio and RMS Exit Speed), showed no significant effects between higher or lower percentages of light aircraft. Thus, it is concluded that an approximately 10% difference in percentages of light aircraft is not sufficient to result in any significant improvement or deterioration in subject's performances.

Subject's Control Activity

Another subject's behavior related aspect, which is subject's control activity, was also investigated. Again, based on different traffic mixes, subject's behavior based on control activity showed insignificant value with $p > 0.05$.

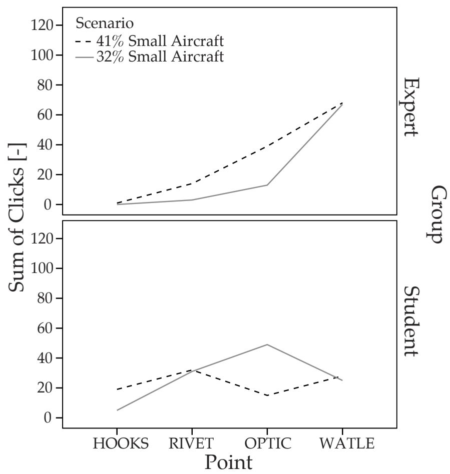


Figure 4.15: Direct to's activity based on traffic mixes.

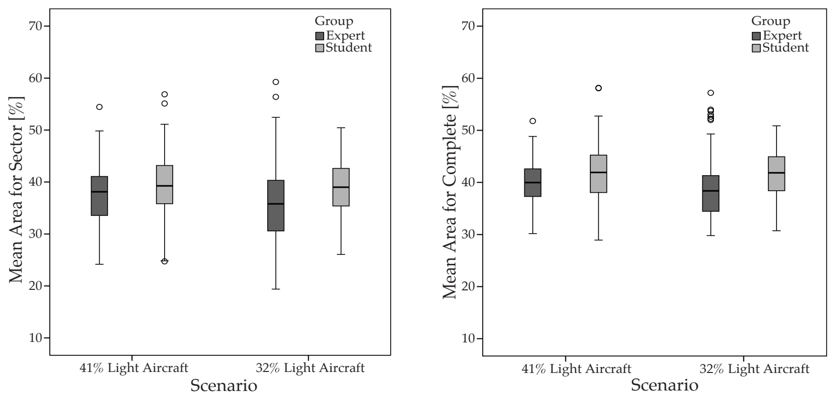
However, differences in the way a command is given in relation to direct to's can still be observed based on different group as well as traffic mix in Figure 4.15. As observed in incoming aircraft proximity and number of streams cases, expert group preferred the final route point as merging point while student group prefer either the second (RIVET) or third (OPTIC) route point as merging point.

Solution Space Diagram Area Analysis

The SSD area properties of different traffic mix was investigated using Wilcoxon-Signed Rank test. Based on the results, only the expert group showed a significantly different SSD Complete Area properties as a result of different traffic mixes, with $Z = -2.201$, $p = 0.028$. The complete result of workload rating can be observed from Table 4.11.

Table 4.11: Wilcoxon-Signed Rank test on the SSD for Sector and Complete Area based on traffic mixes.

	Sector Area		Complete Area	
	Student	Expert	Student	Expert
Z	-0.943	-2.201	-0.943	-1.782
p	0.345	**0.028**	0.345	0.075



(a) Sector Area **(b)** Complete Area

Figure 4.16: SSD area properties based on traffic mixes.

Figure 4.16 illustrates the SSD area properties for Sector Area (Figure 4.16a) and Complete Area (Figure 4.16b) area representation. Based on the figures, it is observed that a slightly higher SSD area properties were found in cases where more light aircraft were present. However, comparison cannot be made due to the fact that for workload rating, both controller groups showed no significant results whereas for the SSD area properties, only the expert group showed significant results. Moreover, the significance was not maintained in the post-hoc analysis with a Bonferroni significance cutoff at 0.025.

4.3.4 Intercept Angle

Finally, the effect of intercept angle on workload was investigated with the exploration of taskload rating, performance metrics and control activity behavior for different intercept angles. In this section, two intercept angles are being investigated namely, 45° and 100° intercept angle. The average workload rating and SSD area properties from the 20 minutes experiment duration was used in the analysis.

Workload rating

In this experiment, angle difference of 55° was chosen. Analysis using Wilcoxon-Signed Rank test has showed that no significant effect was found in the workload rating for both group with Z = -0.524, p = 0.600 (expert) and Z = -0.943, p = 0.345 (student). Figure 4.17 illustrates the workload rating behavior based on different intercept angle settings. Both group of controllers showed the same workload rating behavior in both intercept angle situations but with a larger workload rating dispersion in 45° intercept angle for student group.

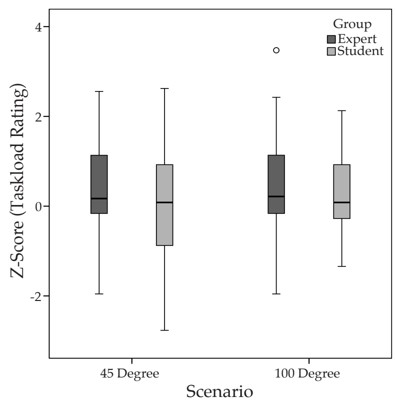


Figure 4.17: Workload rating based on intercept angle.

Subject's Performance

In analyzing the behavior of subjects, several performance metrics were investigated based on different intercept angle properties. Based on the result in Table 4.12, all three performance metrics resulted in a significant effect namely, the RMS Exit Speed, the number of aircraft delivered and RMS Extra Distance Ratio. For the number of aircraft delivered, only the expert group showed significant difference, whereas for RMS Extra Distance and RMS Exit Speed, only the student group showed significant difference. Figure 4.18 illustrated the performance metrics for different intercept angle cases.

Table 4.12: Wilcoxon-Signed Rank test on performances based on intercept angle.

	RMS Exit Speed		Number of Aircraft Delivered		RMS Extra Distance Ratio	
	Student	Expert	Student	Expert	Student	Expert
Z	-2.201	-0.943	-1.134	-2.121	-1.992	-0.734
p	**0.028**	-0.345	0.257	**0.034**	**0.046**	0.463

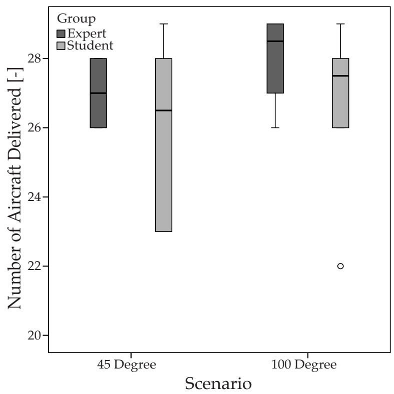


Figure 4.18: Number of aircraft delivered based on intercept angle.

Closer inspection on Figure 4.18 revealed that more aircraft were delivered with 100° intercept angle if compared to 45° intercept angle. One possible explanation is that the 100° intercept angle sector complexity variable result in a shorter distance than the case with 45° intercept angle.

As for RMS Extra Distance and RMS Exit Speed, it is discovered that 100° intercept angle case result in a larger RMS Exit Speed and also RMS Extra Distance Ratio (Figure 4.19a and 4.19b). Larger values of RMS Extra Distance Ratio found in

100° intercept angle might be due to rerouting process of aircraft coming from the direction of 100° intercept angle in order to avoid congestion at merging point.

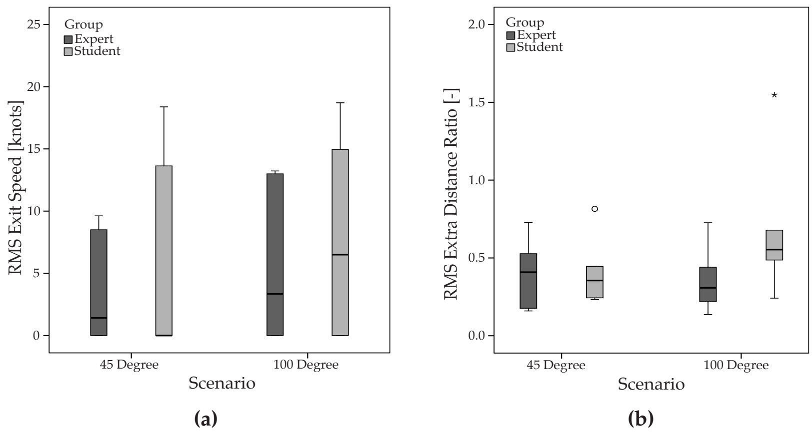
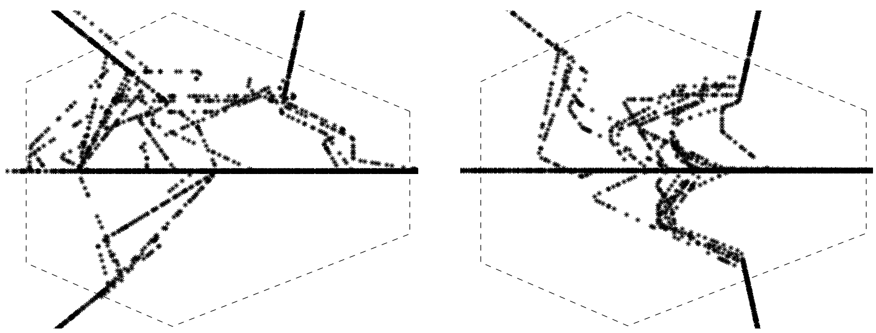


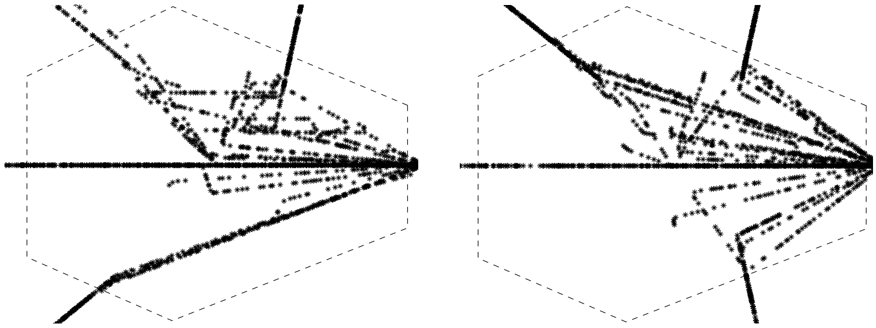
Figure 4.19: Subject's performance based on intercept angle. (a) RMS Exit Speed. (b) RMS Extra Distance Ratio.

An example of traffic management of a student and an expert can be seen in Figure 4.20 and 4.21 where it can be observed that in 100° intercept angle the student tend to reroute aircraft in the opposite direction of the final route point, thus increasing the RMS Extra Distance Ratio. This has also resulted in smaller differences in the number of delivered aircraft for student group. Also, a larger RMS Exit Speed is found in 100° intercept angle believed to be the consequence to maintained good separation between aircraft at the exit point. More aircraft have to be accommodated on the final route stretch as the student continues to merge aircraft with the route through the middle route point.



(a) A student at 45° **(b)** A student at 100°

Figure 4.20: Example of traffic management pattern of a single student.



(a) An expert at 45°
(b) An expert at 100°

Figure 4.21: Example of traffic management pattern of a single expert.

However, post-hoc analysis by applying a Bonferroni correction significance cutoff at 0.025 revealed that the effect seen in performance metrics based on different intercept angle in a sector is insignificant.

Subject's Control Activity

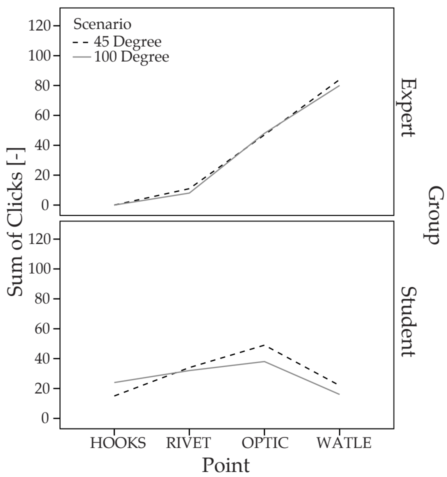


Figure 4.22: Direct to's activity based on intercept angle.

Wilcoxon-Signed Rank test on subject's control activity revealed that there is no significant difference between $N_{Command}$ and N_{Air} for different intercept angle. However, differences between group's direct to's command are similar to previous cases where expert prefer directing incoming aircraft towards the final route point (Figure 4.21a and 4.21b) whereas student prefer to merge aircraft with the route through the middle two route point (Figure 4.20a and 4.20b) before exiting the sector. The summary of sum of clicks based on route points is illustrated in Figure 4.22.

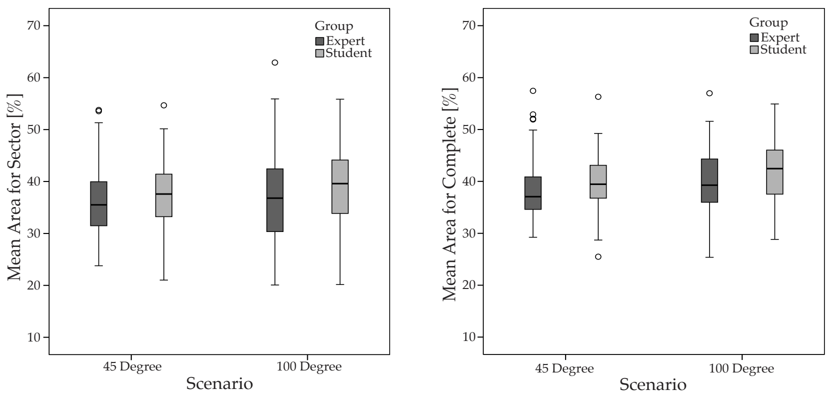
Solution Space Diagram Area Analysis

Analysis on the SSD area properties on different intercept angle revealed that no significant difference were found in both Complete Area and Sector Area for both student and expert groups. Table 4.13 summarize the overall results.

Table 4.13: Wilcoxon-Signed Rank test on the SSD for Sector and Complete Area based on intercept angle.

	Sector Area		Complete Area	
	Student	Expert	Student	Expert
Z	-0.943	-0.743	-1.153	-1.363
p	0.345	0.463	0.249	0.173

Figure 4.23 is illustrated in order to observed SSD area properties behavior pattern. Based on Figure 4.23 it is observed that the SSD area properties for 100° intercept angle situation is larger than 45° intercept angle situation. However, difference of 55° angle between the two scenarios have failed to reveal any significant results in both workload rating and SSD area properties. As discussed in the subject's performance section, subject's control strategy might have altered the initially designed 100° intercept angle, and in the end, influenced the complexity construct of the situation. Thus, conclusion on whether a smaller or larger intercept angle would induce a higher or smaller workload cannot be drawn.



(a) Sector Area **(b)** Compete Area

Figure 4.23: SSD area properties based on intercept angle.

4.3.5 Correlation Analyses

Correlation analyses were conducted to determined how well the SSD area properties correlate with a number of independent and dependent variables, namely the subject's workload ratings, the N_{All}, N_{Sect}, $N_{Command}$, and lastly, N_{Air}. The correlation study were conducted based on the overall situation where all subjects were grouped as one and also based on the grouping of expert and student, to investigate the possible effect of experienced and inexperienced subjects. The analysis was conducted using the Kendall's tau correlation analysis (test statistic R) based on data gathered during a 20 minutes experiment run.

In the following subsection, first the correlation between subject's workload rating will be investigated in relation to SSD area properties. Secondly, the correlations between subject's SSD area properties and a number of dependent and independent measures, namely control activity ($N_{Command}$ and N_{Air}). Thirdly correlation between subject's SSD area properties and the number of aircraft present on the screen (N_{All}) or in the sector (N_{Sect}) are analyzed. Finally, the correlation between workload rating and the latter dependent and independent measures will be investigated.

Correlation between SSD Area Properties and workload rating

The correlation between four SSD area properties and workload rating were investigated and the results are presented in Table 4.14. Based on results, all SSD area properties showed significance value of $p < 0.001$ with SSD Complete Area showing the largest correlation with R = 0.256 (all), R = 0.219 (student) and R = 0.304 (expert). The results highlight that the Complete Area relates to the workload experienced by controllers better than the Sector Area. This concludes that controllers observed not only aircraft within the assigned sector, but also future incoming traffics.

Table 4.14: Correlation coefficients between SSD area properties and workload rating.

Group	Sector Area	Complete Area
All	0.248	**0.256**
Student	0.206	**0.219**
Expert	0.296	**0.304**

Figure 4.24 shows the plots of workload rating in regards to the SSD area properties taken every minute in all six scenarios (6 scenarios x 20 ratings per scenario = 120 rating sequence) for one subject from the group expert. The plots illustrate the level of correlation between the parameters as gathered from the analysis.

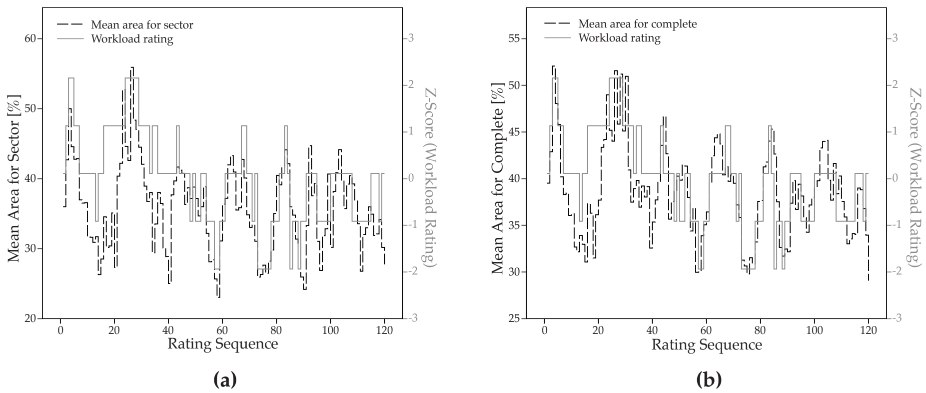


Figure 4.24: Plots of parameter comparison for one subject. (a) SSD Sector Area to workload rating (b) SSD Complete Area to workload rating.

Correlation between SSD Area Properties and Dependent Measures

Cross correlation analysis was performed between SSD area properties and control activity ($N_{Command}$ and N_{Air}). Results are shown in Table 4.15. Cross correlation with control activity revealed that all combination showed significance value of $p < 0.001$. The results can be observed from Table 4.15 with largest correlation value highlighted. Based on the results, SSD Complete Area showed the largest correlation with R = 0.223 (all), R = 0.211 (student) and R = 0.145 (expert) for $N_{Command}$ and R = 0.219 (all), R = 0.182 (student) and R = 0.198 (expert) for N_{Air}. This concludes that the Complete Area is important in evaluating the effort needed to control the traffic.

Table 4.15: Correlation coefficients between SSD area properties and control activity.

Group	$N_{Command}$		N_{Air}	
	Sector Area	Complete Area	Sector Area	Complete Area
All	0.206	**0.223**	0.207	**0.219**
Student	0.203	**0.211**	0.161	**0.182**
Expert	0.135	**0.145**	0.201	**0.198**

Figure 4.25 shows the plots of $N_{Command}$ and N_{Air} in regards to the SSD Complete Area for the 120 rating sequence based on one single subject from the expert group (the same subject).

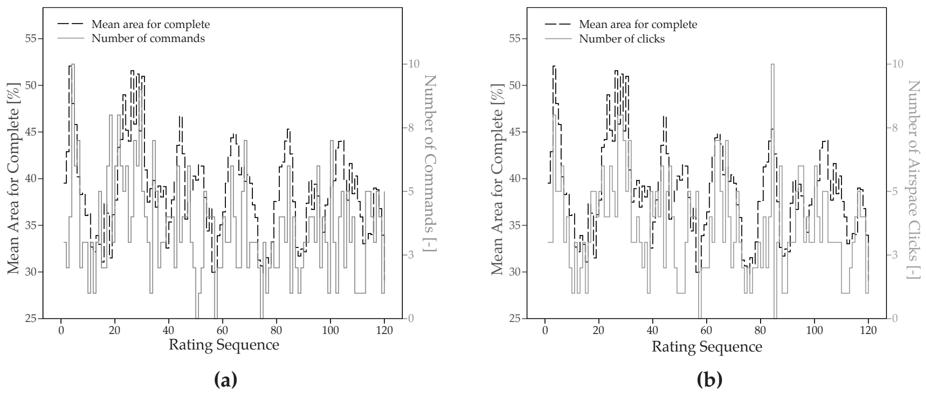


Figure 4.25: Plots of parameter comparison for one subject. (a) SSD Complete Area to $N_{Command}$ (b) SSD Complete Area to N_{Air}.

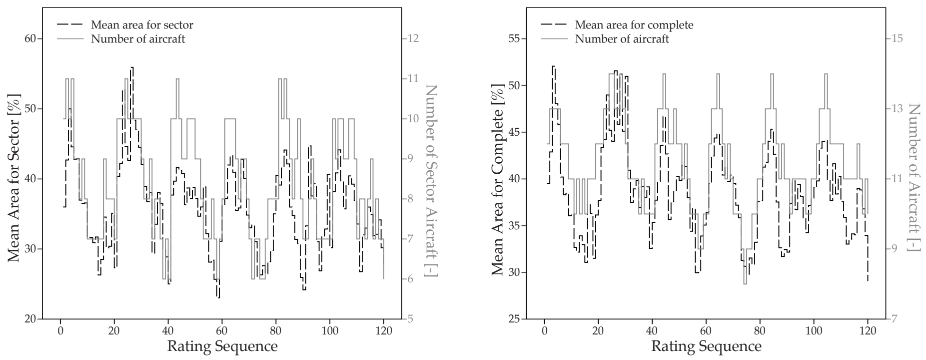
Correlation between SSD Area Properties and Number of Aircraft

As the number of aircraft has been a known complexity measure, the investigation on the correlation between SSD area properties and the number of aircraft presented to the subject is important. Results of the analysis between are laid out in Table 4.16 with all combination showing a significance value of $p < 0.001$. Based on the results, the SSD for Sector Area have the largest correlation with N_{Sect} with R = 0.552 (all), and the SSD for Complete Area have the largest correlation with N_{All} with R = 0.592 (all). In Table 4.16 with the largest correlations are highlighted.

Table 4.16: Correlation coefficients between SSD area properties and number of aircraft.

Group	N_{Sect}		N_{All}	
	Sector Area	Complete Area	Sector Area	Complete Area
All	**0.552**	0.532	0.538	**0.592**
Student	0.484	**0.514**	0.466	**0.559**
Expert	**0.598**	0.512	0.584	**0.592**

Figure 4.26 shows the plots of N_{All} and N_{Sect} in regards to the SSD area properties of Complete Area for the 120 rating sequence based on one single subject from the expert group (the same subject).



(a) (b)

Figure 4.26: Plots of parameter comparison for one subject. (a) SSD Sector Area to N_{Sect}. (b) SSD Complete Area to N_{All}.

Correlation between Number of Aircraft and Independent and Dependent Measures

N_{Sect} and also N_{All} towards other dependent measures were also investigated and the result is laid out in Table 4.17. It is observed that the number of aircraft have the largest correlation with N_{Air} with R = 0.246 (all). This is consistent with the fact that the number of aircraft relate best to how busy an ATCO is. It is also observed that in the student group, the number of aircraft have the largest correlation with $N_{Command}$ (R = 0.260) whereas for expert group the largest correlation was observed with N_{Air} (R = 0.184). This highlights the difference in the way both groups behave in the presence of traffic. The student group respond to the number of aircraft by giving commands, whereas the expert group respond by monitoring the traffic (by clicking on aircraft).

Table 4.17: Correlation coefficients between number of aircraft and dependent measures.

Group	N_{Sect}		N_{All}	
	$N_{Command}$	N_{Air}	$N_{Command}$	N_{Air}
All	0.215	0.236	0.223	**0.246**
Student	**0.260**	0.237	0.250	0.241
Expert	0.087	0.170	0.085	**0.184**

Correlation between Workload Rating and Independent Variables and Dependent Measures

Cross correlation between workload ratings and the latter independent and dependent measures were conducted and the results are laid out in Table 4.18. All of the results gathered shows a significance value of $p < 0.001$ with workload rating having the largest correlation with N_{Air} with $R = 0.198$ (all), $R = 0.298$ (expert). However, in student's group case, the workload rating showed the largest correlation with the N_{All} with $R = 0.165$.

Table 4.18: Correlation coefficients between workload rating and independent variables and dependent measures.

Group	n_{Sect}	n_{All}	$N_{Command}$	N_{Air}
All	0.174	0.191	0.194	**0.198**
Student	0.135	**0.165**	0.162	0.140
Expert	0.226	0.231	0.280	**0.298**

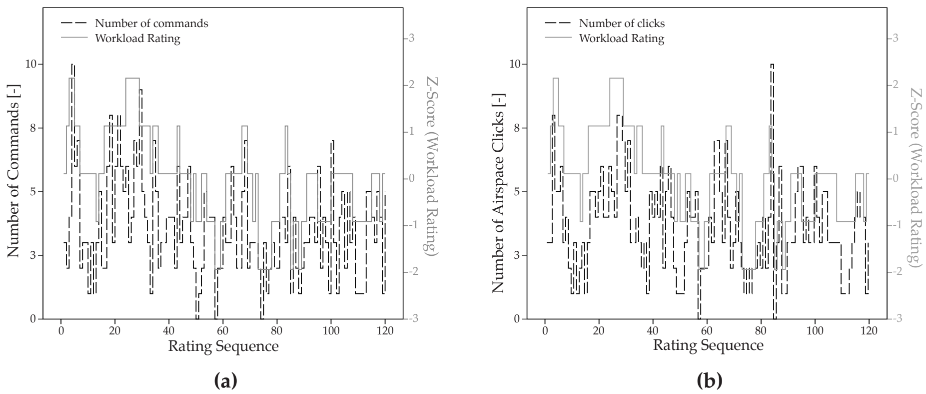
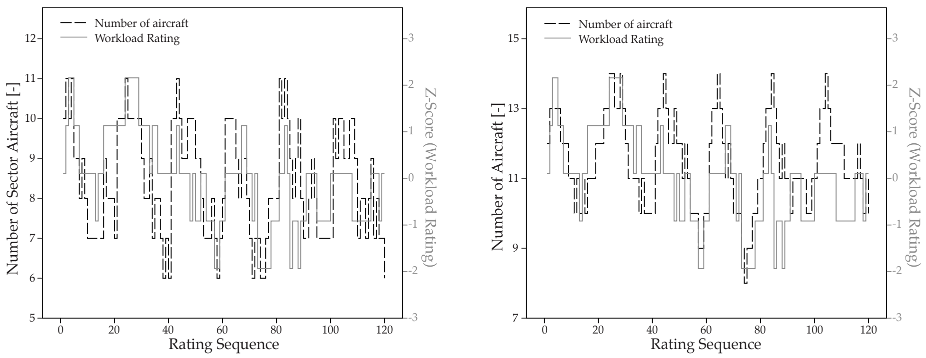


Figure 4.27: Plots of parameter comparison for one subject. (a) $N_{Command}$ to workload rating. (b) N_{Air} to workload rating.

This also highlights that the workload matches the activity needed for the expert group and the number of aircraft for the student group. Figure 4.27 and 4.28 shows the plots of $N_{Command}$, N_{Air}, N_{Sect} and N_{All} in regards to the workload rating for the 120 rating sequence based on one single subject from the expert group (the same subject).



(a) **(b)**

Figure 4.28: Plots of parameter comparison for one subject. (a) N_{Sect} to workload rating. (b) N_{All} to workload rating.

4.4 Discussion

This chapter investigated the effect of several sector design variables towards workload and SSD metric. It is gathered that the experiment was unable to confirm the findings on the theoretical study conducted previously (Chapter 3) on the effects of traffic complexity variables on the SSD. In each sector complexity factor exploratory analysis, significant results were gathered in either workload ratings or SSD area properties, but never in both. Thus, comparison between workload and SSD area properties cannot be made. It should also be highlighted that in dynamic situations, sector complexity is highly shaped by human control behavior, and sector complexity issues are not simply about how aircraft affect one another, but more towards on how an action affect the complexity of the sector.

In the case of incoming aircraft proximity, higher incoming aircraft proximities resulted in higher SSD area properties. However, conclusion on whether higher incoming aircraft proximity would induce higher workload cannot be drawn as the two different horizontal proximities scenarios did not result in a significant difference between workload ratings. Having said that, control activity was also seen higher with high incoming aircraft proximity with a larger $N_{Command}$ and N_{Air} in high incoming aircraft proximity situation. However, in terms of subject performance, no significant effect was found between high and low incoming aircraft proximity situations. Thus, even when the control behavior is different, subjects managed to maintain the same performance metrics for both scenarios.

The workload ratings based on different numbers of streams showed higher workload ratings when more streams were present in both cases for the expert group.

However, for SSD area properties, no significant result was observed. Analysis on the performance metric revealed that a different control behavior pattern was observed in the student group when different percentages of light aircraft were present (Case 2) together with the varying number of streams. Differences in the way the student group handled traffic in Case 2 reveals that differences in scenario designs results in a different controller strategy. However, the difference in control strategy was not sufficient enough to trigger a significant change in the perception of taskload.

Difference in traffic mix did not affect the workload ratings of both controller groups. On the other hand, the SSD area properties were significantly different only in the expert group. Higher SSD area properties gathered from the expert group were observed when traffic has more light aircraft. However, comparison cannot be made, as the approximately 10% difference in light aircraft percentage present within the sector did not result in any significant workload rating differences.

The assignment of 55° in intercept angle difference in this experiment also appeared to be insufficient to demonstrate possible differences in workload ratings and also SSD area properties. This is may be due to differences in managing incoming aircraft strategy (Figure 4.20 and 4.21) that has altered the initially designed 100° intercept angle, and thus influenced the complexity construct of the situation. On the other hand, subject's performance showed significant results in the Number of Aircraft Delivered (expert group) and RMS Extra Distance Ratio (student group) and RMS Exit Speed (student group) with consistently larger values in 100° intercept angle scenarios.

For all the scenarios representing different complexity factor situations, both group of controllers showed a different pattern of control activity. It is observed that the expert group consistently preferred the final route point as merging point while student group prefer either the second (RIVET) or third (OPTIC) route point as merging point.

However, it must be mentioned here that separate statistical analyses were performed for each individual complexity factor. Although reckoned with in the analyses by adopting appropriate correction factors, possible interaction effects were neither exposed nor taken into account. For example, differences in the percentages of light and heavy aircraft in combination with variations in incoming traffic sequences could have tainted the obtained results. As such, the results obtained in this chapter could potentially sketch an overly favorable or unfavorable picture of some of the complexity factors in terms of their correlations to workload.

Results in the previous sections have indicated certain level of correlations between SSD area properties, workload rating and also other dependent and independent measures. Based on the results, the workload rating showed the largest correlation towards the SSD area properties with R = 0.256 (Complete Area) and R = 0.248 (Sector Area) compared to other sector complexity measure where R = 0.174 (N_{Sect}), R = 0.191 (n_{All}), R = 0.194 ($N_{Command}$) and R = 0.198 (N_{Air}). The SSD area properties also showed a good correlation with N_{Sect} having R = 0.532 (Complete Area) and 0.552 (Sector Area) and with N_{All} having R = 0.592 (Complete Area) and 0.538 (Sector Area).

4.5 Conclusion

In each sector complexity factor investigation, significant results were gathered in either workload rating or SSD area properties, but never in both. Thus, the hypothesis regarding horizontal proximities, number of streams, traffic mixes and intercept angle cannot be confirmed. A small sample size and scenario design which may represent an insufficient difference between two different level of sector complexity might have hampered the possibility of investigating the effect of different sector complexity design towards controller's workload rating and SSD area properties.

It is also concluded that constructing two different level of single sector complexity variable were found to be the biggest challenge for this experiment. It is almost impossible to change one sector complexity variable, without affecting another sector complexity variable. For example, when adding another stream of aircraft, it would affect (1) the proximity of traffic within the sector, (2) the additional intercept angle with the route that the additional stream creates, (3) the dispersion of traffic within the sector and the chain effect goes on. This situation would occur when trying to implement changes in any other sector complexity variable. Thus, a more controlled experiment environment in order to really understand the effect of single sector complexity variable towards controller workload is needed.

However, despite the fact that both groups performed differently and had different control strategies, based on the correlation analyses, the SSD area properties have the best correlation to the workload as indicated by the subject, especially in case where the Complete Area projected to the controller is taken into consideration, showing that the SSD could be a good workload predictor.

Regardless of this positive result, it should be used with caution due to the limited sample size used in the experiment and the individual variations in reporting workload and control behavior. Although we realize that we cannot generalize the experimental findings to the whole population, it does show that investigating sector complexity as a single factor within a dynamic environment is very difficult to do.

5

Solution Space in Conflict Detection Scenarios

This chapter continues the investigation on the use of Solution Space Diagram (SSD) as a measure of sector complexity and also a predictor of performance and workload, focusing on scenarios regarding an Air Traffic Controller (ATCO)'s ability to detect future conflicts. A human-in-the-loop experiment in which two sector complexity variables were varied, namely traffic density and intercept angle has been designed and conducted. The experiment tested varying degrees of intercept angle and two levels of traffic density within the same sector layout. A short experiment duration and a single predetermined conflict in each scenario, ensure a controlled experimental environment. The main aim of the experiment is to investigate whether the SSD can predict the workload ratings and subject performance in a conflict detection task. The experiment also explores whether there is a common threshold in the SSD area properties where controllers would begin to detect conflict pair. As expected, the experiment results revealed that higher traffic density generates higher workload and SSD area properties. On the other hand, no common pattern can be observed, which can directly associate workload rating and SSD area properties in a varying intercept angles situation. As conflict presented in the experiment were between converging aircraft, smaller SSD observation angle were found to correlate better with the workload rating. These results were anticipated, as in converging conditions aircraft ahead of the velocity vector will capture the main focus. The SSD also does not represent a trigger for conflict detection with no consistent SSD area percentage prediction to represent the time ATCO would start to detect conflict.

5.1 Introduction

Controller's workload and sector complexity has been an important topic of research in the Air Traffic Control (ATC). Based on previous research projects, the ability of the controller to ascertain whether or not an aircraft pair will lose separation (more commonly known as conflict detection), is affected by variety of variables that include but not limited to sector properties such as convergence angle (Remington et al., 2000, Rantanen & Nunes, 2005, Nunes & Kirlik, 2005, Lee et al., 2009) and number of aircraft (Endsley & Rodgers, 1998).

In Chapter 3, a preliminary investigation into the two-aircraft situation with direct collision path, while have varying intercept angle were discussed. Results from the preliminary investigation (Figure 5.1) has shown that a larger Solution Space Diagram (SSD) area percentage is gathered with smaller Time To Conflict (TTC). It is also concluded that smaller intercept angles produce a larger SSD area percentage for the same TTC.

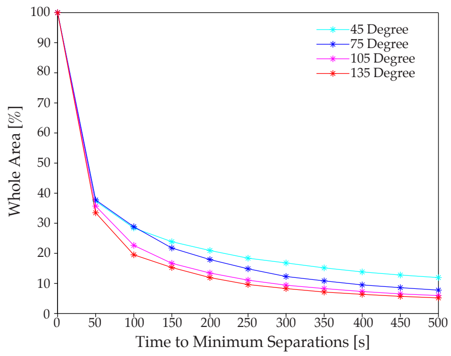


Figure 5.1: SSD area progression with Time To Conflict (TTC) for different intercept angle.

This experiment is aimed at systematically analyzing the capability of the SSD to illustrate changes in the sector complexity, and also at using the method to obtain an objective measurement of the sectors complexity. The experiment is also intended to investigate whether conflict detection time correlates to the size of the covered SSD area.

40 scenarios, each lasting two minutes with a single predetermined conflict were investigated in the experiment. Scenarios contained either 8 or 14 aircraft and with the conflict at one of the five different intercept angles, namely 30°, 60°, 90°, 120° and 150°, were explored. All of the scenarios were presented in the same sector layout. First, the effect of traffic density towards workload and SSD area will

be explored, followed by the effect of varying intercept angle based on different traffic density level. Lastly, to explore whether the SSD can be a measure of sector complexity, the correlation between SSD area properties and workload rating and subject's performances will be investigated.

Shorter runs were used in for this experiment to enable more scenario repetitions, providing more data. The setup is also intended to allow a more controlled environment, which ensures that controller behavior has limited influenced on the traffic in the sector.

5.2 Experimental Design

5.2.1 Subjects and Tasks

A total of 10 male subjects participated in the study. The test group subject represents a population of subjects that participated in an extensive ATC introductory course or have extensive knowledge of Air Traffic Controller (ATCO) tasks. The subjects' age ranged between 27 and 50 years (μ = 33.10, σ = 7.89). The subjects were instructed to identify and resolve a future separation violation problem in a two-minutes scenario situation.

The experiment was conducted using the standalone simulator illustrated in Figure 5.2. The display consists of two parts, the Plan View Display (PVD) area and the control panel area. The left part of the screen is the PVD area that shows the sector under control, the surrounding area of this sector and the aircraft within the area.

Color coding was used for the aircraft symbols. When in conflict, the aircraft involved would turn red. A green tag would mean that the aircraft is flying free and is not in conflict with other aircraft. One aircraft could be selected at a time. The selected aircraft would also show a circle around it, representing the 5 NM separation radius. Aircraft symbols were shown with a tag, and the information in the tag included callsign, current and intended flight level, current speed, intended heading and the type of aircraft.

The right part of the display contains the control panel area where heading command can be given to selected aircraft. The subject can only give commands to aircraft that are inside the controlled area. The outer and inner circles on the command display correspond to the maximum and minimum speed of the selected aircraft, respectively. A thick line represents the current speed and heading vector of the aircraft. To instruct an aircraft to change it's heading, the controller can position the mouse to the desired heading within the maximum and minimum speed circles. A single mouse click on the desired heading will instruct the aircraft to change its

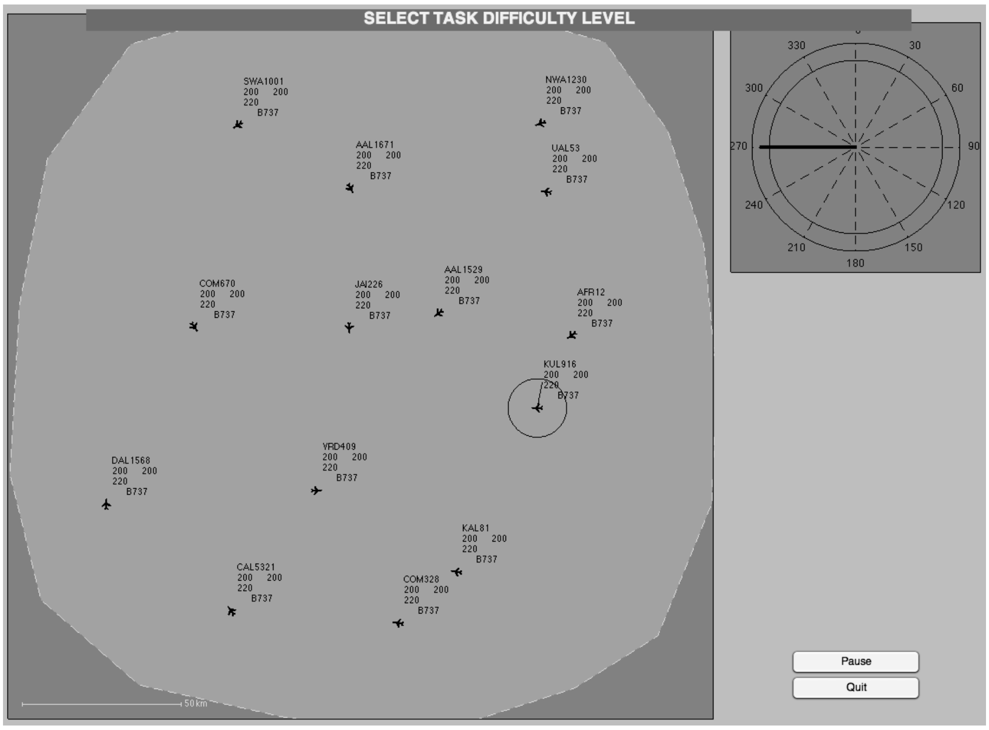


Figure 5.2: Experiment simulator.

heading.

The subjects were instructed to detect a pair of aircraft that was on a collision course as quickly as possible. They then had to try to resolve the future conflict by giving only heading instructions, after which they had to direct the aircraft to its original heading again. The conflicting pair will experience a direct collision within the next 120 seconds and there is one type of aircraft present with a known speed limit of 200 - 240 knots.

During the experiment, the participants were asked to rate their perceived workload every 30 seconds. An automated stimulus provided a message on the display (Figure 5.2) that triggered the participants to rate their workload by means typing a number between 1 (low workload) and 7 (high workload) on the keyboard.

5.2.2 Independent Variables

The independent variables in the experiment are: (1) the intercept angle of the conflict pair, which has five levels: 30°, 60°, 90°, 120° and 150°, and (2) the traffic density, which has two levels: low (8 aircraft) and high (14 aircraft) traffic density. The independent variables provide a total number of (5 x 2 =) 10 experiment conditions (Table 5.1).

Table 5.1: Scenario settings.

Scenario	Angle Setting	Number of Aircraft
1	30°	14
2	60°	14
3	90°	14
4	120°	14
5	150°	14
6	30°	8
7	60°	8
8	90°	8
9	120°	8
10	150°	8

Each traffic density level has five background scenarios. This is done with the intention to gather more data, through enabling more scenario runs, while making sure that subjects do not recognize the conflict pair. The 40 scenarios sequence for each subject are shown in Table 5.2.

Table 5.2: Sequence of scenario based on subject.

	High Traffic					Low Traffic					High Traffic					Low Traffic				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Subject 1	A4	B3	C2	D1	E5	K1	F5	G4	H3	J2	D3	E2	A1	B5	C4	H5	J4	K3	F2	G1
Subject 2	B5	C4	D3	E2	A1	F2	G1	H5	J4	K3	E4	A3	B2	C1	D5	J1	K5	F4	G3	H2
Subject 3	C1	D5	E4	A3	B2	G3	H2	J1	K5	F4	A5	B4	C3	D2	E1	K2	F1	G5	H4	J3
Subject 4	D2	E1	A5	B4	C3	H4	J3	K2	F1	G5	B1	C5	D4	E3	A2	F3	G2	H1	J5	K4
Subject 5	E3	A2	B1	C5	D4	J5	K4	F3	G2	H1	C2	D1	E5	A4	B3	G4	H3	J2	K1	F5
Subject 6	A3	B2	C1	D5	E4	K5	F4	G3	H2	J1	D2	E1	A5	B4	C3	H4	J3	K2	F1	G5
Subject 7	B4	C3	D2	E1	A5	F1	G5	H4	J3	K2	E3	A2	B1	C5	D4	J5	K4	F3	G2	H1
Subject 8	C5	D4	E3	A2	B1	G2	H1	J5	K4	F3	A4	B3	C2	D1	E5	K1	F5	G4	H3	J2
Subject 9	D1	E5	A4	B3	C2	H3	J2	K1	F5	G4	B5	C4	D3	E2	A1	F2	G1	H5	J4	K3
Subject 10	E2	A1	B5	C4	D3	J4	K3	F2	G1	H5	C1	D5	E4	A3	B2	G3	H2	J1	K5	F4
	High Traffic					Low Traffic					High Traffic					Low Traffic				
	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Subject 1	D1	E5	A4	B3	C2	F5	G4	H3	J2	K1	A1	B5	C4	D3	E2	H5	J4	K3	F2	G1
Subject 2	C4	D3	E2	A1	B5	K3	F2	G1	H5	J4	E4	A3	B2	C1	D5	G3	H2	J1	K5	F4
Subject 3	B2	C1	D5	E4	A3	J1	K5	F4	G3	H2	D2	E1	A5	B4	C3	F1	G5	H4	J3	K2
Subject 4	A5	B4	C3	D2	E1	H4	J3	K2	F1	G5	C5	D4	E3	A2	B1	K4	F3	G2	H1	J5
Subject 5	E3	A2	B1	C5	D4	G2	H1	J5	K4	F3	B3	C2	D1	E5	A4	J2	K1	F5	G4	H3
Subject 6	D5	E4	A3	B2	C1	F4	G3	H2	J1	K5	A5	B4	C3	D2	E1	H4	J3	K2	F1	G5
Subject 7	C3	D2	E1	A5	B4	K2	F1	G5	H4	J3	E3	A2	B1	C5	D4	G2	H1	J5	K4	F3
Subject 8	B1	C5	D4	E3	A2	J5	K4	F3	G2	H1	D1	E5	A4	B3	C2	F5	G4	H3	J2	K1
Subject 9	A4	B3	C2	D1	E5	H3	J2	K1	F5	G4	C4	D3	E2	A1	B5	K3	F2	G1	H5	J4
Subject 10	E2	A1	B5	C4	D3	G1	H5	J4	K3	F2	B2	C1	D5	E4	A3	J1	K5	F4	G3	H2

For high traffic density, the background scenarios are indicated in Table 5.2 as A, B, C, D and E and for low traffic density, the background scenarios are indicated in Table 5.2 as F, G, H, J and K. The presentation order of the first 20 scenarios was randomized (in a batch of 5 scenarios) to counterbalance a possible order effect on the dependent measures in the experiment. The last 20 scenarios were the representation of the first 20 scenarios, which was rotated at 180° and again randomize in a batch of 5 scenarios.

5.2.3 Dependent Measures

The dependent measures in the experiment consisted of performance behavior, number of aircraft properties, SSD area properties, and subjective workload ratings.

i) Performances behavior: Two subjects' performance measures were evaluated in this experiment, which is the time measures and also the conflict measures. In the time measures, three performance instances were looked into, namely the time from starting of the scenario to identifying the conflict pair ($T_{identify}$), the time it takes to think of a proper solution (T_{think}) and the time from starting of the scenario to resolution of the conflict pair ($T_{resolve}$). Figure 5.3 summarizes the main performances measures. All of the time measures were gathered using the history of controller's control activity. For example, the $T_{identify}$ is measured based on the time controller click on the right conflict pair and is not followed by any action of selecting other aircraft than the conflict pair before starting resolution act.

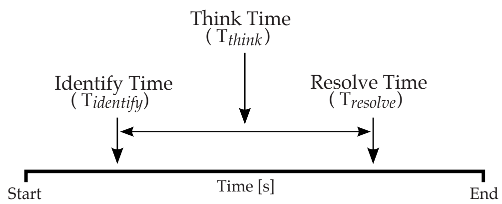


Figure 5.3: Time performance measures.

There are three types (Type 1, Type 2 and Type 3) of subject-induced conflict measures ($N_{Conflict}$) gathered at the end of each scenario run. These conflicts arise as a result of controller action and it is measured on all aircraft within the sector. Type 1 represent long term conflict that will occur between the next 5 to 10 minutes, Type 2 represent medium term conflict that will occur between the next 2 to 5 minutes and Type 3 represent short term conflict that will occur in less than 2 minutes. These conflicts are gathered based on all aircraft present in the sector. Since each scenario only have one predetermined conflict, any other conflict which arises after the resolution action is considered to be a subject-induced conflict.

ii) Subject's workload rating: The workload rating, measured on a one to seven scale, were provided by the subject every 30 seconds during the experiment run. In order to correct for inter-subject differences, Z-scores of the subjective ratings were used in the subsequent data exploration. This correction was performed by calculating the Z-scores for every test subject.

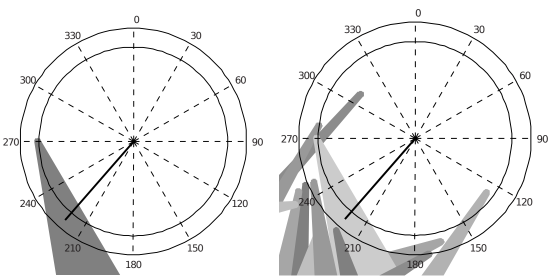
iii) SSD area properties: Three SSD area properties was measured, namely the Conflict Area ($A_{conflict}$), the Mean of Total Area (A_{total}) and the Mean Area (A_{mean}). The SSD area properties were measured every 30 seconds (during the workload rating instances) and also based on the time measures (during $T_{identify}$ and $T_{resolve}$ instances). The $A_{conflict}$ represent unsafe area caused only by the conflicting aircraft ($A_{converge}$) and the A_{total} represent the mean unsafe area (A_{whole}) of the two conflicting aircraft. The A_{mean}, on the other hand, represent the sum of unsafe area A_{whole} for all individual aircraft in the sector divided by the total number of aircraft in the sector. All measures are gathered based the equation as follow:

$$A_{conflict_{i,j}} = \frac{A_{converge_{i,j}} + A_{converge_{j,i}}}{2}$$

$$A_{total_{i,j}} = \frac{A_{whole_i} + A_{whole_j}}{2}$$

$$A_{mean} = \frac{1}{n}\sum_{i=1}^{n} A_{whole_i}$$

The A_{whole} is calculated using the total area covered within minimum and maximum velocity-heading band (aircraft performance limit) of each individual aircraft, while $A_{converge}$ is calculated using the total area covered within the same aircraft performance limit but focusing only towards the two conflicting pair. Examples of the SSD showing the $A_{converge}$ and A_{whole} situations are illustrated in Figure 5.4.



(a) $A_{converge}$ **(b)** A_{whole}

Figure 5.4: Example of SSD with either one or multiple observed aircraft.

Previous research by d'Engelbronner et al. (2010) found that in a merging scenario, a strong trend towards higher correlations for larger observation angles were observed. In this experiment, analysis based on different observation angles were conducted in order to see if the limit of controller observation boundary will changed, when the task of identifying conflict pair were given to a controller. Several observation angle conditions were measured in this experiment. The observation angle was defined as the semi-sided angle relative to the velocity vector (Figure 5.5). Based on these settings, observation angle of either 45° (OB_{45}), 90° (OB_{90}) or 180° (OB_{180}) were assigned to the aircraft heading. As a result, a total of nine main SSD area properties (3 SSD area x 3 observation angle = 9 SSD) were investigated in this experiment.

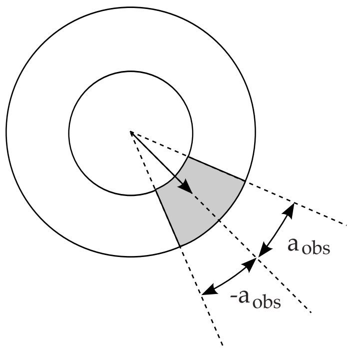


Figure 5.5: SSD observation angle.

5.2.4 Procedure

On every scenario subjects were instructed to detect a pair of aircraft that was on a collision course as quickly as possible. They are then given the task to resolve the future separation violation by giving heading instructions, after which they had to direct the aircraft to its original heading again. The procedure of the experiment consists of having the subject run through 48 scenarios including eight training runs. The experiment was run at 4 times real-time, similar to previous experiment procedures.

The results of the training sessions will not be used in the analysis and are considered as a learning phase in which the subject could get accustomed to the control actions needed for the experiment. The last 40 sessions are considered to be the measurement sessions of which the recorded data will be used in the analysis. A single run lasts approximately 2 minutes.

5.2.5 Hypotheses

The experiment was intended to study the effect of two sector complexity variables, namely intercept angle and traffic density in a two-minutes scenarios towards controller workload and SSD area properties. It is hypothesized that the SSD properties are related to both controllers workload and sector complexity. It is also hypothesized that a certain configuration of these complexity variables will lead to different effects on the controller's performances, workload and also SSD properties. The hypotheses concerning complexity metric are as follow:

1. **Intercept Angle:** It is hypothesized that a direct collision course with the same TTC, a larger intercept angle will result in lower workload rating and also SSD area properties (Figure 5.1). This can be due to the fact that the distance between aircraft is larger for larger intercept angle if the same TTC value is to be maintained.
2. **Traffic Density:** It is hypothesized that higher traffic density result in higher workload rating and also SSD area properties. This is based on the notion that the more aircraft present within a sector, the more effort is needed to carefully scrutinize possible conflict pair.
3. **SSD as measure of sector complexity and also a predictor of performance and workload:** It is hypothesized that the SSD area properties can correctly measure sector complexity and also predict performance and workload in a conflict detection task. Higher workload and lower performance (for example late conflict detection or resolution) should also reveal a higher SSD area properties. It is also hypothesized in this chapter that a lower workload and SSD area properties relate to a lower sector complexity.

5.3 Results

In this section, the effect of sector complexity variable changes to workload, controller performances and SSD area properties is discussed, followed by a correlation analysis between the measures. Finally, conclusions on the behavior of traffic density and intercept angle towards workload and also SSD are drawn. In this result section, effects were considered significant at a probability level $p \leq 0.05$, where p is the probability that the null hypothesis is true.

Background Scenario Effect

In the experiment, five different background scenarios were used for each level of traffic density. Before making the assumption that the data can be treated as one, and ignoring differences in other traffic present in the scenarios, analysis on the initial A_{mean} of all possible scenarios and the reported controller workload ratings were conducted.

A one-way ANOVA (test statistic F) was used to test whether the initial A_{mean} of possible background scenario and intercept angle combinations significantly different from one another. Based on the analysis, there is no significant different between the background scenario's A_{mean}, with $F(4, 20) = 0.063$, $p = 0.992$ and $F(4, 20) = 0.153$, $p = 0.959$ for high and low traffic and high traffic scenarios, respectively. This has shown that the multiple background scenarios initial settings, do have the same level of complexity.

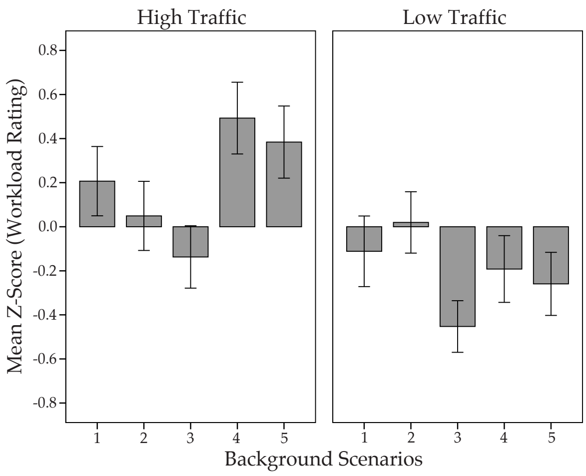


Figure 5.6: Mean of workload rating based on different background scenarios.

Wilcoxon-Signed Rank analysis (test statistic Z) revealed that 7 out of 20 combinations of background scenarios showed significant differences between one another (with $p < 0.05$) , based on the reported workload ratings. These background scenarios combinations are gathered based on 10 combinations of background scenarios observed through five different intercept angles on two traffic density levels. Figure 5.6 illustrates the workload rating of different background scenario grouped according to different traffic density level. The variation in workload rating within single background scenario is the result of variation in intercept angles. However, with a Bonferroni correction cutoff at 0.01, these significant results were no longer valid. Thus, it is concluded that the effects of having different background scenarios can be fully neglected in this experiment.

5.3.1 Traffic Density

Firstly, workload rating, SSD area properties and subjects performance are considered in elaborating the effect of traffic density. Each variable was investigated by means of Wilcoxon-Signed Rank test to explore the influence of each intercept angle towards subjects and SSD. The workload, time measures and SSD area measures were analyzed using the average data gathered from 20 two-minutes experiment runs based on different scenario settings. This is done in order to gather an overall controller workload situation and SSD area properties of each sector complexity factor.

Workload Rating

Wilcoxon-Signed Ranks test revealed that as expected, the different levels of traffic density significantly affects workload ratings, with Z = -2.701, p = 0.007. A higher traffic density resulted in a higher workload rating by the subject. This can be observed in Figure 5.7. This result matches the notion that more aircraft result in a more demanding work environment.

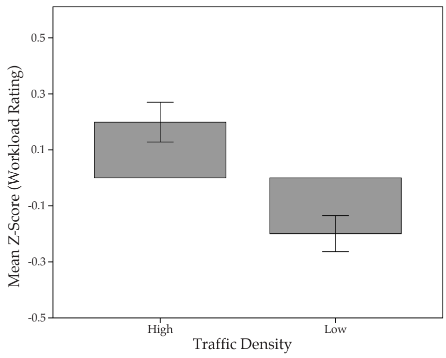


Figure 5.7: Mean of workload rating based on traffic density.

Time Measures

The subject's time measure performance behavior corresponds with the workload rated by the controller. Based on the results, the level of traffic density significantly affect the $T_{resolve}$, with Z = -2.191, p = 0.028. The $T_{identify}$ and T_{think} are, however, not significantly affected by different levels of traffic density, with Z = -1.580, p = 0.114 and Z = -0.051, p = 0.959, respectively. Figure 5.8 illustrate the effect of traffic density towards $T_{identify}$ and $T_{resolve}$.

Both $T_{identify}$ and $T_{resolve}$ performance measures showed a higher sum of ranks in favor of higher traffic density level. This also means that a higher $T_{identify}$ and $T_{resolve}$ is needed to resolve conflicts relating to a higher traffic density. However, the same could not be confirmed after post-hoc analysis by applying a Bonferroni correction significance cutoff at 0.025.

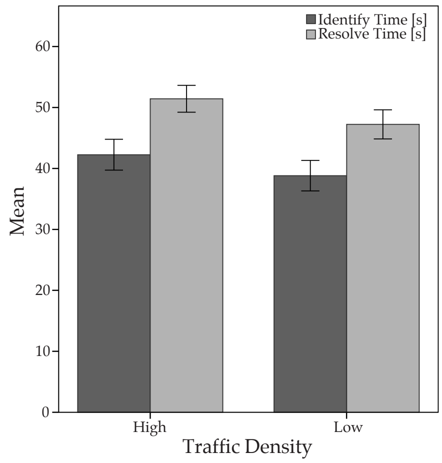


Figure 5.8: Mean of $T_{identify}$ and $T_{resolve}$ based on traffic density.

Subject-induced Conflicts

Different types of conflict represent different time in the future where they will occur. Type 1 represent conflicts that will occur between the next 5 to 10 minutes, Type 2 represent conflict that will occur between the next 2 to 5 minutes and Type 3 represent conflict that will occur within the next 2 minutes.

In an overall analysis, it is gathered that a total of 120 scenarios out of 400 scenarios ended up with controller-induced conflicts as a result of controller's control actions. A Pearson Chi-Square test (test statistic χ^2) was performed to determined if future separation violation were distributed differently across different traffic density. The test indicate a significant different in the possibility of the end scenario resulting in conflict, with $\chi^2(1) = 6.857$, $p = 0.009$. 28% of low traffic scenarios and 12% of high traffic scenarios, end with controller-induced conflicts. The result also highlight that 60% out of 120 cases of subject-induced conflict come from low traffic density situations. This can be observed from Figure 5.9.

The investigation on future separation violation is extended by classifying the types of conflict based on different traffic density. Based on the analysis, it is observed that difference in traffic density has an effect on the types of conflict, with $\chi^2(1) = 28.654$, $p < 0.001$. A surprisingly high number of Type 3 conflicts are observed at the end of low density scenario runs, even when controller performance and workload were recorded and reported to be lower in these cases.

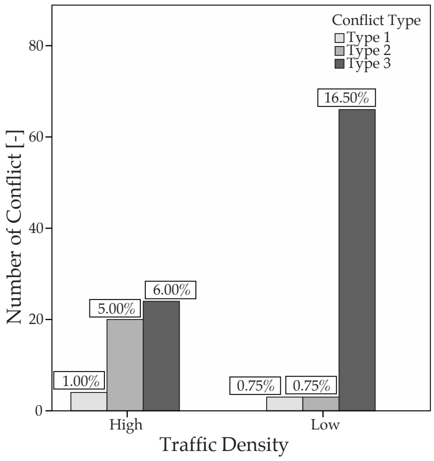


Figure 5.9: Conflict type based on traffic density.

Solution Space Diagram Area Properties

The SSD area properties were initially investigated based on the level of traffic density to see whether indeed the SSD can depict the effect of number of aircraft. Table 5.3 showed the results of Wilcoxon-Signed Rank analysis on the SSD area properties. Based on the results, it is observed that several SSD measures showed to be significantly affected by the different level of traffic density.

Table 5.3: Wilcoxon-Signed Rank test on the SSD for three different observation angles based on traffic density.

	OB_{180}			OB_{90}			OB_{45}		
	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}
Z	-1.274	-1.274	-2.803	-2.191	-1.988	-2.803	-2.803	-1.784	-2.803
p	0.203	0.203	**0.005**	**0.028**	**0.047**	**0.005**	**0.005**	0.074	**0.005**

Figure 5.10 illustrates the A_{mean} properties of three different observation angles grouped according to different levels of traffic density. It is observed that a higher

traffic density results in a higher SSD area properties. This behavior corresponds with the behavior gathered on the controller's workload rating and also performances. It is also observed that A_{mean} in OB_{180} case showed a larger difference between the two traffic densities compared to OB_{45} cases.

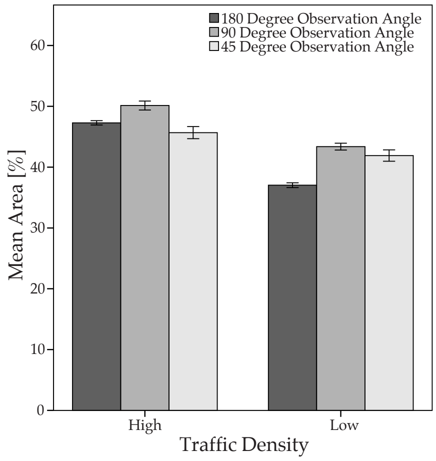


Figure 5.10: A_{mean} area properties based on traffic density.

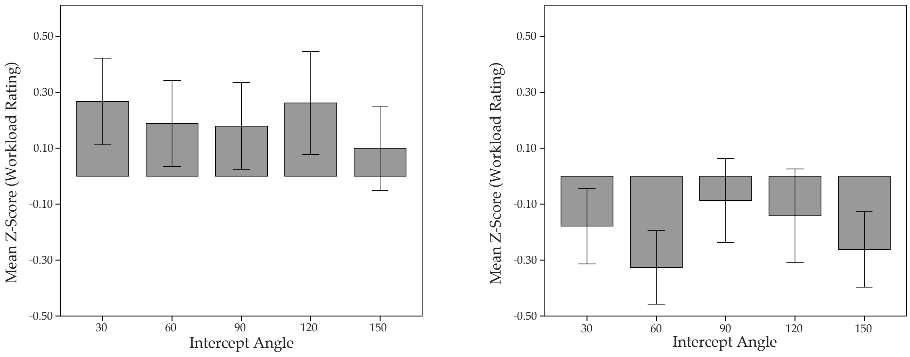
5.3.2 Intercept Angle

Secondly, effects of intercept angle on subject's workload rating, SSD area properties and controller's performance were investigated. Each variable was investigated by means of Friedman analysis to look into the overall effect of intercept angle and later elaborated with a Wilcoxon-Signed Rank test to explore the influence of each intercept angle. Here, the workload, time measures and SSD area measures were also analyzed using the average data gathered from eight two-minutes experiment runs based on different intercept angle settings per traffic density.

Workload Rating

Based on Friedman analysis (test statistic χ^2), both high and low traffic density situations of eight two-minutes experiment run showed no significant effect on workload rating in varying intercept angle settings, with $\chi^2 = 5.705$, $p = 0.222$ and $\chi^2 = 3.155$, $p = 0.532$, respectively. It is believed that changes in constructing varying intercept angle settings might not be significant enough when combined with the background scenarios.

Figure 5.11 illustrate plots of mean of workload rating in respect to different intercept angle. It can be observed that there are two sets of decreasing workload rating for both high and low traffic density situation. For high traffic density, the workload rating is decreasing between 30° to 90° intercept angle, and also between 120° to 150° intercept angle. For low traffic density, a similar pattern is visible but between 30° to 60° intercept angle, and also between 90° to 150° intercept angle. However, differences in the workload rating for various intercept angle is not significant enough to be captured in the statistical analysis.



(a) High Traffic **(b)** Low Traffic

Figure 5.11: Mean of workload rating based on intercept angle.

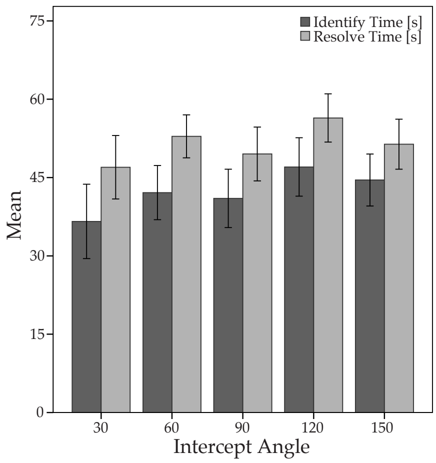
Time Measures

Friedman analysis was also conducted to assess the effects on controller's time performance behavior namely, $T_{identify}$, $T_{resolve}$ and T_{think}. The results of the analysis are shown in Table 5.4. It is observed that significant difference as a result of varying intercept angle is only present in $T_{resolve}$ and only in high traffic density. However, the $T_{identify}$ and T_{think} are not significantly affected by the differences in intercept angle for both high and low traffic density situations. This shows that despite differences in the $T_{resolve}$, subjects showed no significant difference in the time they identified a conflict pair and also the time they needed to figure out a proper solution for separation violation problems with different intercept angles.

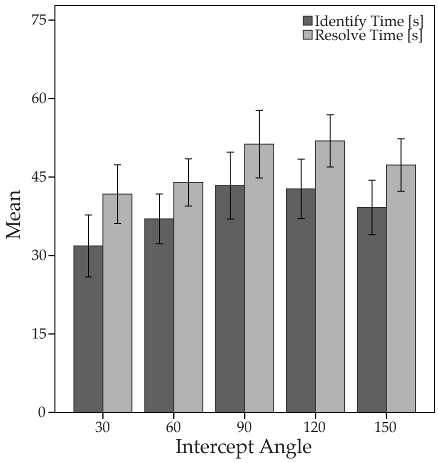
$T_{identify}$**:** Figure 5.12 illustrates the fluctuation of $T_{identify}$ for different intercept angles in both high and low traffic density situations. The $T_{identify}$ is also shown to be divided into two sections as observed in workload rating. For high traffic density situation, two sets of increasing and decreasing $T_{identify}$ can be observed. While in low traffic density situation, the $T_{identify}$ is observed to be increasing from 30° to 90° and then it starts decreasing again.

Table 5.4: Friedman's ANOVA test on controller's performance measures based on intercept angle.

	$T_{identify}$		$T_{resolve}$		T_{think}	
	High Traffic	Low Traffic	High Traffic	Low Traffic	High Traffic	Low Traffic
χ^2	8.640	5.360	12.880	8.080	7.440	2.240
p	0.071	0.252	**0.012**	0.089	0.114	0.692



(a) High Traffic



(b) Low Traffic

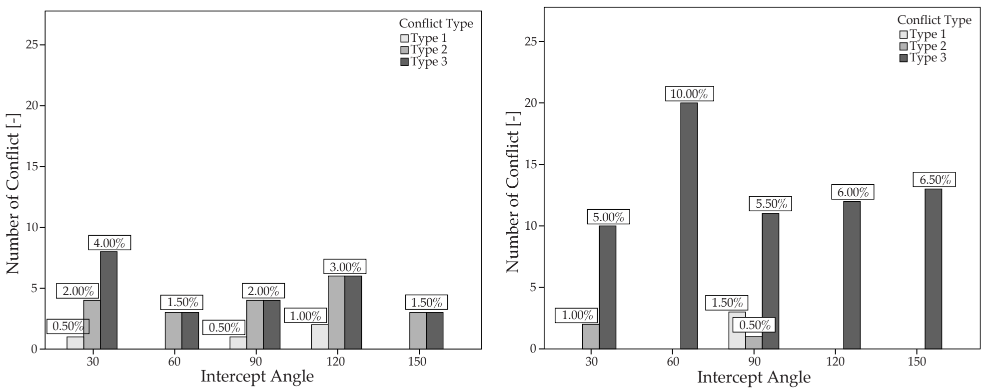
Figure 5.12: Mean of $T_{identify}$ and $T_{resolve}$ based on intercept angle.

$T_{resolve}$**:** A post-hoc analysis using Wilcoxon-Signed Rank test with a Bonferroni correction (p $<$ 0.01) for high traffic density revealed that only the $T_{resolve}$ of the 90° intercept angle is significantly different from the $T_{resolve}$ of the 120° intercept angle, with Z = -2.701, p = 0.007. It is observed from Figure 5.12 that the $T_{resolve}$ for the 90° intercept angle has a lower mean rank than that of the 120° intercept angle. This also means that lesser time is needed to resolve conflict in the 90° intercept angle cases compared to the 120° intercept angle cases.

In low traffic density, the $T_{resolve}$ showed no significant effect towards differences in intercept angle (with a Bonferroni correction of $p < 0.01$). Figure 5.12 illustrates the effect of intercept angle differences towards $T_{resolve}$ with the same fluctuation pattern in $T_{identify}$ for high traffic density situation. For low traffic density situation, the same pattern of increasing and decreasing $T_{resolve}$ were found with a peak at 90° intercept angle.

Subject-induced Conflicts

The analysis did not gather any significant effects of different intercept angle on the possibility to end with subject-induced conflict for both high traffic and low traffic density situations with $\chi^2(4) = 7.840$, $p = 0.098$ and $\chi^2(4) = 4.905$, $p = 0.297$, respectively. Figure 5.13 illustrates the number of scenarios that ends with a future separation violation based on varying traffic density and intercept angle. The low traffic density situation was shown to have a higher number of future separation violation compared to high traffic density situation. This also shows that even with an earlier detection of conflict in low traffic density situation and also a lower workload rating, it does not ensure a good overall awareness of the consequences of subject's action.



(a) High Traffic **(b)** Low Traffic

Figure 5.13: Conflict type based on intercept angle.

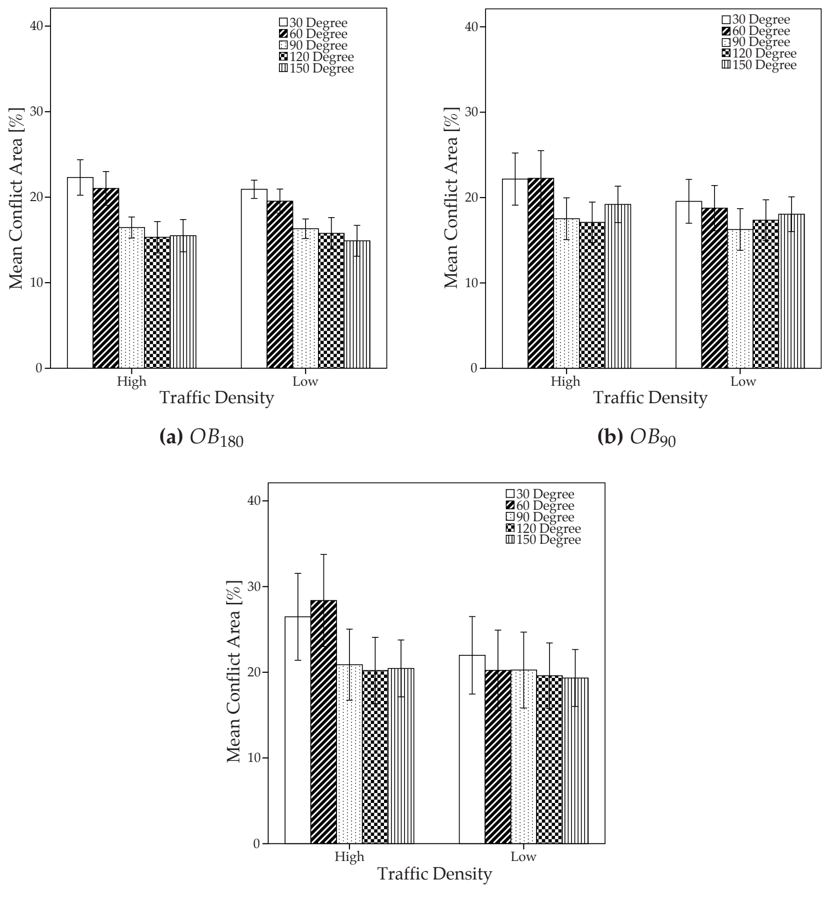
The analysis also gathered no significant influence of different intercept angles on the type of future separation violation in high traffic density situation, with $\chi^2(8) = 2.821$, $p = 0.945$. However in low traffic density situation, different intercept angle significantly affect the type of subject-induced conflicts with $\chi^2(8) = 18.982$, $p = 0.015$. These can be observed from Figure 5.13b where a significantly larger number of Type 3 conflicts were present at the end of the scenario.

Solution Space Diagram Area Properties

Friedman analysis gathered a number of $A_{conflict}$ area properties that showed significant effect towards intercept angle. This can be observed from Table 5.5. Further analyses have gathered a number of intercept angle combinations that resulted in a highly significant difference in the SSD area properties. However, there are no apparent patterns of intercept angle combinations that would lead to a highly significant difference in the SSD area properties. The behavior of the SSD area towards differences in intercept angle will be further discussed using Figure 5.14.

Table 5.5: Friedman's ANOVA test on SSD area properties based on intercept angle.

		OB_{180}			OB_{90}			OB_{45}		
		A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}
High	χ^2	2.240	30.400	0.720	6.800	18.160	0.080	7.360	17.360	0.560
	p	0.692	<**0.001**	0.949	0.147	**0.001**	0.999	0.118	**0.002**	0.967
Low	χ^2	0.320	28.850	2.960	5.200	6.560	3.440	1.200	0.400	1.680
	p	0.988	<**0.001**	0.565	0.267	0.161	0.487	0.878	0.982	0.794



(a) OB_{180} **(b)** OB_{90} **(c)** OB_{45}

Figure 5.14: $A_{conflict}$ area properties based on intercept angle.

Figure 5.14 showed the significantly affected SSD area properties ($A_{conflict}$) for three different observation angles. For high traffic density situation, it is observed that the $A_{conflict}$ is decreasing from 30° until 120° intercept angle. Then, it is followed by a small increase in the $A_{conflict}$ value. This can be seen for both OB_{180} and OB_{90}. The decreasing and increasing $A_{conflict}$ values is observed to be in accordance to the time performance measures behavior. A relatively higher $A_{conflict}$ values found at 45° intercept angle (Figure 5.14c) is observed to correspond to an earlier conflict identification and resolution, $T_{identify}$ and $T_{resolve}$ (Figure 5.12).

For low traffic density situation, the $A_{conflict}$ showed almost a constant decrease for OB_{180} and OB_{45}. While for OB_{90}, the $A_{conflict}$ showed a decreasing trend followed by an increasing values. However, the differences is not significant as in the high traffic density.

5.3.3 Correlation Analyses

The previous analyses focused on the individual effects of traffic density and intercept angle on controller's workload rating, performance and also SSD area properties. In this section, the correlation between workload rating and performance with SSD area properties were investigated. The analysis was conducted using the Kendall's tau correlation analysis (test statistic R) based on data gathered during a 40 two-minutes experiment run in both high and low traffic density.

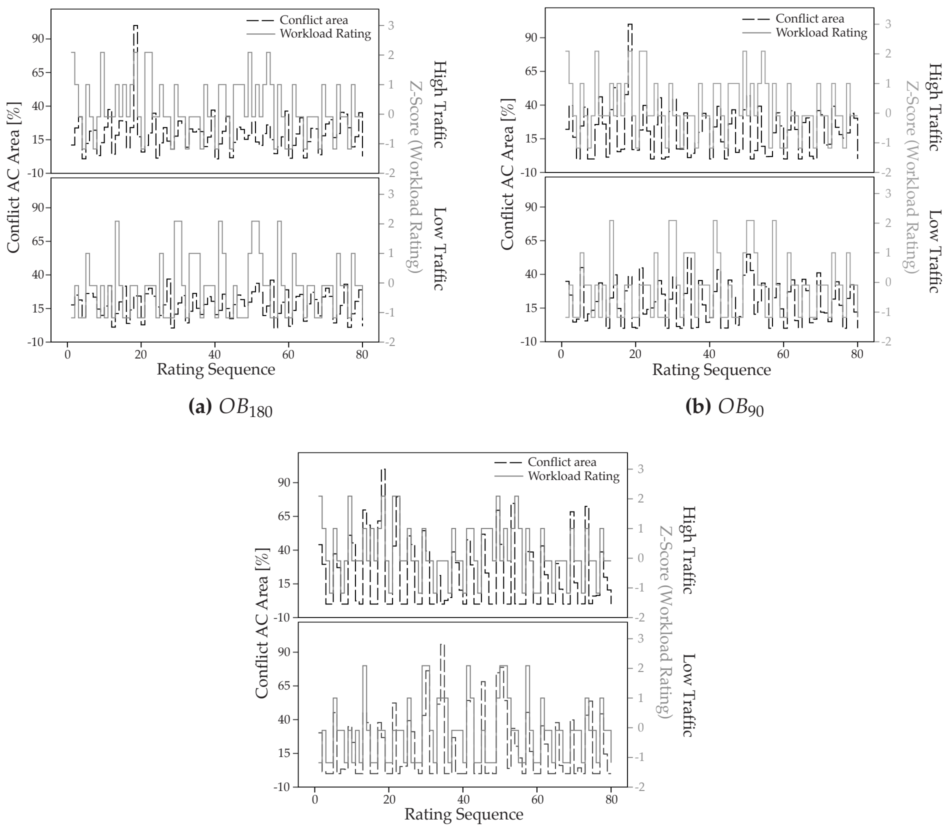
Correlation between Workload rating and SSD area properties

The correlation between workload rating and SSD area properties were investigated and the results are shown in Table 5.6. The area properties are investigated based on three observation angles, which are OB_{45}, OB_{90} and OB_{180}. Based on the results, all combination showed significance value of $p < 0.001$ except for three instances in OB_{180} (marked with a '*').

Table 5.6: Correlation coefficient between workload rating and SSD area properties.

Group	OB_{180}			OB_{90}			OB_{45}		
	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}
Overall	0.064	0.055	-0.020*	0.343	0.330	0.296	0.382	**0.427**	0.336
High Traffic	0.010*	0.040*	-0.129	0.320	0.312	0.270	0.357	**0.409**	0.289
Low Traffic	0.107	0.075	-0.243	0.368	0.352	0.268	0.399	**0.447**	0.372

*Correlation at higher than 0.05



(a) OB_{180} **(b)** OB_{90} **(c)** OB_{45}

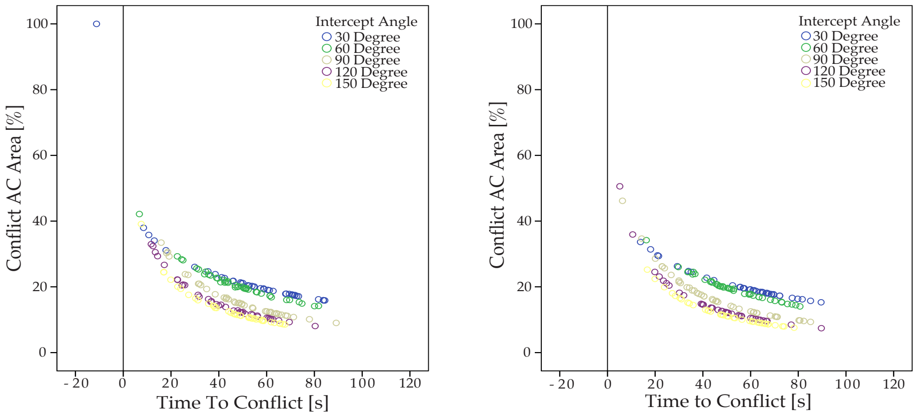
Figure 5.15: Workload rating with $A_{conflict}$ at three different observation angles.

It is also observed from Table 5.6 that the largest correlations were detected between workload rating and $A_{conflict}$ using OB_{45} (highlighted in Table 5.6 in bold). This also indicated that the area within the OB_{45} is best at representing the controller's workload rating compared to other area properties. The fact that SSD area properties of OB_{45} have a better correlation with workload rating suggests that the area which is in the direction of the velocity vector has more impact in determining the level of difficulty that subject's has to undergo in a scenario where one separation violation situation is known to take place.

Figure 5.15 illustrate the correlation between the $A_{conflict}$ area property at OB_{180}, OB_{90} and OB_{45} with workload rating. Based on the figure, it is visible that the $A_{conflict}$ of OB_{45} does have the highest correlation with the workload rating. The figure illustrated are based on the same subject participated in the experiment, which is illustrated in the previous section.

Correlation between SSD Area Properties and Time Measures

To compare with the results gathered from initial analysis of two aircraft with future direct collision path in Chapter 3 (also illustrated in Figure 5.1), analysis on SSD area properties with regards to time measures ($T_{identify}$ and $T_{resolve}$) were conducted. Figure 5.16 illustrates the scatter plots of $A_{conflict}$ for OB_{180} gathered from the experiment. The plot matches the outcome of the initial analysis of two aircraft with future direct collision path in Figure 5.1. The same pattern were gathered as the $A_{conflict}$ focuses solely on the two conflicting aircraft, similar to the previous case studies in Chapter 3.

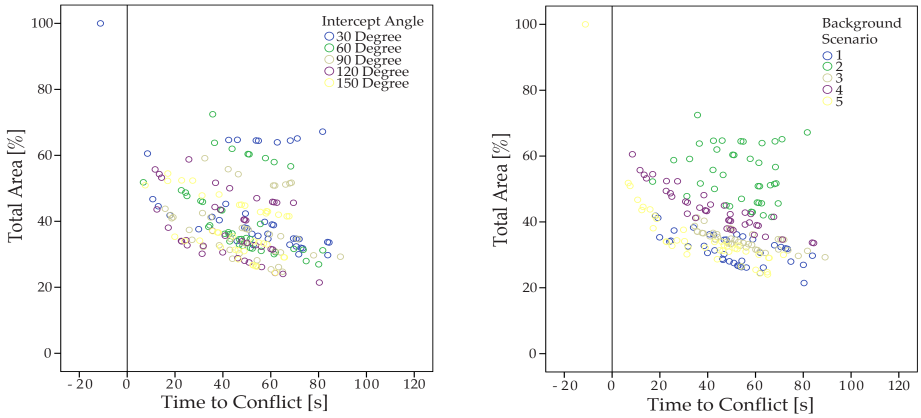


(a) High traffic **(b)** Low traffic

Figure 5.16: $A_{conflict}$ for OB_{180} at the moment of conflict identification ($T_{identify}$) based on intercept angle.

To show the effects of other aircraft within sector, A_{total} for OB_{180} were illustrated and compared with the previous $A_{conflict}$ findings. When considering the A_{total}, the effect of other aircraft within the sector became more predominant (Figure 5.17) as it changes the pattern of SSD area. However, the same relation between intercept angle and the SSD area where smaller intercept angle has higher SSD area properties is visible within the same background scenario. This has shown that the background scenario does has an effect in the behavior of the A_{total}, but to the same degree that the behavior of smaller intercept angle has higher SSD area properties is still visible. This can be observed when comparing Figure 5.17a to Figure 5.17b.

The results of correlation analysis conducted between $T_{identify}$ and SSD area properties are shown in Table 5.7. Based on the results, the $A_{conflict}$ for OB_{90} showed the highest correlation with $T_{identify}$, with $R = 0.508$ (overall). However, when considering high and low traffic density separately, the A_{mean} for OB_{180} showed the



(a) Different intercept angle **(b)** Different background scenario

Figure 5.17: A_{total} for OB_{180} at the moment of conflict identification ($T_{identify}$) for high traffic density situation.

highest correlation with $T_{identify}$, with $R = 0.611$ (high traffic) and $R = 0.697$ (low traffic). This shows that the A_{mean} for OB_{180} relates best with $T_{identify}$ compared to A_{total} and $A_{conflict}$ based on different level of traffic density, whereas $A_{conflict}$ relates best with $T_{identify}$ when considering the situation as a whole.

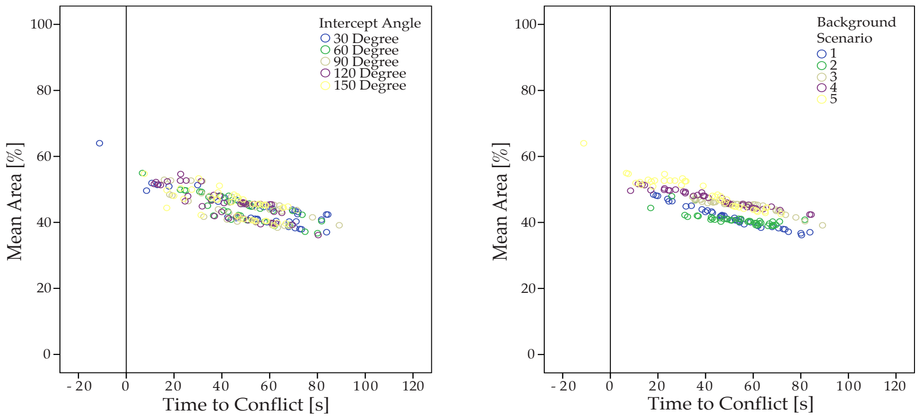
Table 5.7: Correlation coefficient between SSD area properties and $T_{identify}$.

Group	OB_{180}			OB_{90}			OB_{45}		
	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}
Overall	0.351	0.502	0.424	0.307	**0.508**	0.209	0.401	0.505	0.012*
High Traffic	0.251	0.508	**0.611**	0.258	0.506	0.128	0.414	0.513	0.016*
Low Traffic	0.456	0.494	**0.697**	0.364	0.494	0.325	0.394	0.496	-0.032*

*Correlation at higher than 0.05

Figure 5.18 shows the scatter plot of A_{mean} for OB_{180} at the time of conflict identification ($T_{identify}$), in order to illustrates the overall sector complexity construct. Based on the figures, it can be seen that a higher $T_{identify}$ result in a higher A_{mean} for OB_{180} values regardless of the intercept angle (Figure 5.18a) and background scenario (Figure 5.18b). This is expected, due to the fact that the further away in time the situation progresses, the higher the SSD area covered for each individual aircraft, as both sector contain several crossing aircraft. Thus, a much later conflict identification results in a higher SSD area properties.

However, the relation where smaller intercept angle has higher SSD area properties is not visible with the A_{mean}. The A_{mean} of different intercept angles is distributed within the same range (Figure 5.18a), with a strong association to different background scenarios (Figure 5.18b). This also conclude that the effect of background



(a) Different intercept angle **(b)** Different background scenario

Figure 5.18: A_{mean} for OB_{180} at the moment of conflict identification ($T_{identify}$) for high traffic density situation.

scenarios has a more significant impact for A_{mean}, compared to previous situation in $A_{conflict}$ and A_{total} as it overshadows the different intercept angles behavior.

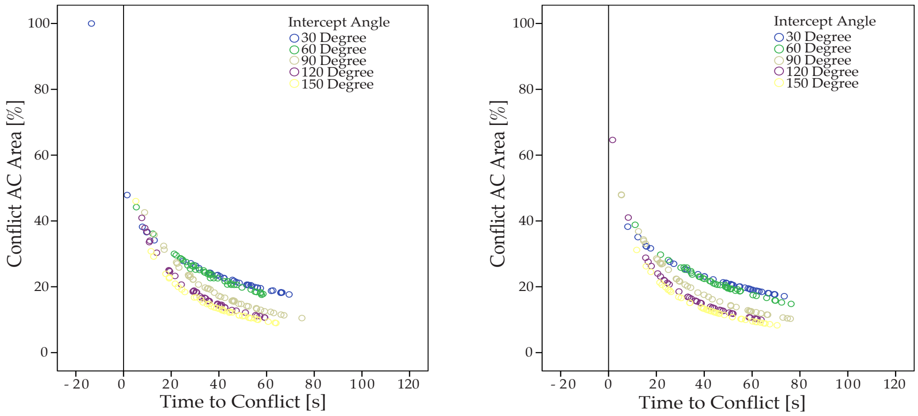
The area percentage for A_{mean} of OB_{180} for both high and low traffic are approximately between 20% to 50% or 40% to 60%, respectively. Even when A_{mean} data showed to be more concentrated than $A_{conflict}$ and A_{total} data, there are still quite a large spread of area percentage covered when the conflict were detected. Thus, no common SSD area properties which may trigger identification of conflict pair was found. It is concluded that the SSD metric is not suitable for prediction of conflict detection time.

Table 5.8: Correlation coefficient between SSD area properties and $T_{resolve}$.

Group	OB_{180}			OB_{90}			OB_{45}		
	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}	A_{total}	$A_{conflict}$	A_{mean}
Overall	0.394	0.572	0.426	0.368	0.568	0.168	0.494	**0.580**	-0.008*
High Traffic	0.288	0.578	0.546	0.320	0.573	0.060**	0.505	**0.582**	-0.021*
Low Traffic	0.499	0.569	**0.607**	0.414	0.566	0.271	0.489	0.575	-0.030*

*Correlation at higher than 0.05

The results of correlation analysis conducted between $T_{resolve}$ and SSD area properties are shown in Table 5.8. Based on the results, the $A_{conflict}$ for OB_{45} showed the highest correlation with $T_{resolve}$, with $R = 0.580$ (overall), $R = 0.582$ (high traffic). The A_{mean} for OB_{180}, however, showed highest correlation with $T_{resolve}$ at low traffic density with $R = 0.607$. The scatter plots of $A_{conflict}$ for OB_{180} at the time of resolution ($T_{resolve}$) is illustrated in Figure 5.19. It is observed that the scatter plots of SSD area properties and $T_{resolve}$ showed the same pattern as gathered in the $T_{identify}$. However, the plots are shifted nearer to the intended time of conflict and also have a slightly higher SSD area properties, as the resolution actions started further in time.



(a) High traffic

(b) Low traffic

Figure 5.19: $A_{conflict}$ at the moment of conflict resolution ($T_{resolve}$) based on intercept angle.

5.4 Discussion

This chapter investigates the use of the SSD a workload predictor in an ATC task of identifying and resolving a single future separation violation. The experiment tested varying degrees of intercept angle and two levels of traffic density within the same sector layout.

Investigating single sector complexity variable in a dynamic environment has shown to be a complicated task. This is due to the fact that the investigation of single sector complexity variable (based on scenario of only two converging aircraft) might not deliver the 'same' effect as it would deliver in 'real' situation. However, adding another element by introducing other non-conflicting aircraft in the sector might interfere with the controller's attention from the issue that is being investigated. A trade-off has to be made between investigating single element of sector complexity variable and presenting a closer to actual condition to ATCO. In this experiment, several background scenarios has been introduced, to present the latter.

To ensure the same level of sector complexity, pre-experiment analysis based on the A_{mean} of all possible background scenarios and intercept angle combinations were conducted. Based on the analysis, it was verified that the scenario and intercept angle combinations were not significantly different between one another. However, post-experiment analysis on the reported workload ratings revealed that 7 out of 20 combinations of background scenarios showed to be significantly different between one another. However, with a Bonferroni correction cutoff at 0.01, these significant results were no longer valid. Thus, it is concluded that the effects of having different background scenarios can be fully neglected in this experiment.

The analysis on the behavior of workload rating and SSD area properties towards changes in two sector complexity variables, namely traffic density and intercept angle was investigated using an average data gathered from 40 two-minutes experiment runs.

Differences in traffic density has resulted in significant difference in workload ratings, with a higher traffic density having higher workload ratings. The effect of different traffic densities was more dominant than other sector complexity variables in the scenario as it involved different levels of effort in controlling and ensuring safe separation between aircraft. As for subject's performance in identifying and resolving conflict and SSD area properties, longer time and higher SSD area percentage were identified for higher traffic density. With these results, we can conclude that a higher traffic density indeed, resulted in a higher sector complexity construct, as expected.

The hypothesis that a larger intercept angle results in a higher workload rating and SSD area properties cannot be fully confirmed. The semi-circular intercept angles investigated in this chapter seems to be divided into two portions. Two sets of decreasing controller's workload rating were observed for both high and low traffic density situations. The first set of decreasing workload rating were observed between 30° to 90° intercept angle for high traffic density situations and between 30° to 60° intercept angle for low traffic density situations. The second set of decreasing workload rating were observed between 120° to 150° intercept angle for high traffic density situations and between 90° to 150° intercept angle for low traffic density situations. The time performance measures ($T_{identify}$ and $T_{resolve}$) and also SSD area properties also showed a similar behavior of the semi-circular intercept angle division, however, with different trends. This could be an indicator of a mental division between acute and obtuse convergence angles. Though, this can only be confirmed with further investigations.

Correlation analysis between SSD area properties and workload rating revealed that $A_{conflict}$ at OB_{45} correlates best with workload rating. Thus, suggesting that the area which are in the direction of the velocity vector has more impact in determining the level of difficulty that subject's has to undergo in a scenario where one separation violation situation is known to take place. Figure 5.20 illustrates the example of the effect of aircraft speed assignment towards controller observation strategy with the gray area representing the observation boundary. It is revealed that when all aircraft within a sector has the same speeds (coupled with single Flight Level (FL) scenario), the problem were construed to be a geometrical problem, thus changing the observation strategy.

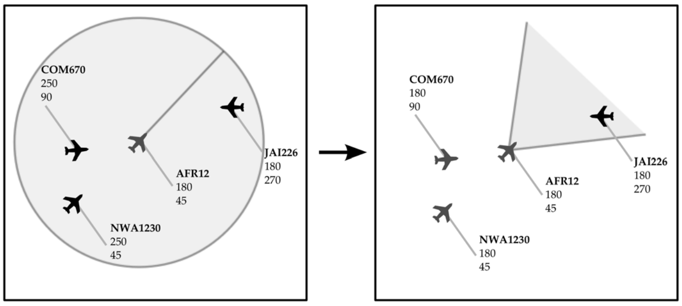


Figure 5.20: Changes in the observation strategy.

In this case, the perception of demand has changed. Each aircraft is no longer observed by looking at other surrounding traffic as a whole, but were only observed by comparing only the converging pairs. This has alter the 'normal' observation strategy from inspecting all aircraft within the vicinity of the controlled aircraft towards only inspecting the ones which intersect in the line of controlled of the aircraft velocity vector.

A larger $A_{conflict}$ based on $T_{identify}$ and $T_{resolve}$ is gathered for smaller intercept angles. This is a result of a larger intercept angle or a bigger horizontal distance between aircraft at identification and resolution instances. However, the sector complexity construct also depends on other aircraft within the sector. These were illustrated through A_{total} and A_{mean}, which also incorporated the surrounding aircraft within the SSD construction. Difference in the behavior of the A_{total} and A_{mean} area properties compared to $A_{conflict}$ indicate that other aircraft within a sector also give an important effect on the space an aircraft has to maneuver.

Nevertheless, the pattern that shorter TTC would result in more area within the SSD to be covered remained the same. This is expected, due to the fact that the further away in time the situation progresses, the higher the SSD area covered for each individual aircraft. Having said that, there are no single area coverage observed when conflict detection usually occurs. Thus, concluding that the SSD metric did not provide a predictor of conflict detection.

Differences in subjects strategies on how to handle a conflict also results in different 'after conflict' effects, which may or may not result in a future separation violation. The fact that 120, out of 400 or 30% of the scenarios end up with future separation issue with 60% of these 120 cases of subject-induced conflicts come from low traffic density situation, also shows that sector complexity is dynamic. It is not only dependent on how a sector is designed, but also highly dependent on the controller's own action.

5.5 Conclusion

The results gathered in this chapter conclude that a higher traffic density indeed, as expected, result in higher workload and SSD area properties. On the other hand, no common pattern can be observed, which can directly associate workload ratings and SSD area properties for various intercept angles. It is concluded, based on the findings in this experiment, that intercept angle is an intricate matter to be investigated as a single sector complexity construct, in a situation where the difficulty of identifying conflicting aircraft pair is not only influenced by controller behavior but also by the neighboring traffic within the sector.

The experiment discovered that a smaller SSD observation angle was found to correlate better with the workload rating, whereas larger observation angles correlate better with controller's performance. This could be the result of the nature of the task that specifically demand subjects to identify and also resolve one conflict pair within the next two minutes. As all aircraft have the same speed, inspection of possible conflict were done mainly between converging aircraft pair and also within a certain radius from of an aircraft. In these conditions, only aircraft ahead of the velocity vector will be considered the main focus. Owing to the fact that the task may have resulted in technical artifact present in the experiment setup, the overall SSD area will be used in the following chapter of comparing the SSD metric to currently available sector complexity measures.

The experiment also did not show a clear threshold on SSD area percentage where a controller would start to detect a conflict pair. Thus, it is concluded that the SSD does not represent a trigger for conflict detection.

6

Sector Complexity Measures: A Comparison

To compare sector complexity measures in terms of their transferability in capturing dynamic complexity across different controllers and sectors, a human-in-the-loop experiment using two distinct sectors has been designed and conducted. Sector complexity measures, such as the intercept angle of traffic routes, the number of crossing points, the clustering of crossing points and entry and exit points, the sector geometry and area varied over the two sectors. The experiment results revealed that the Solution Space Diagram (SSD) metric has a higher correlation with the controllers' workload ratings than the number of aircraft and the unweighted NASA Dynamic Density (DD) metric. Although linear regression analysis improved the correlation between the workload ratings and the weighted DD metric as compared to the SSD metric, the DD metric proved to be more sensitive to changes in sector layout and groups of controllers than the SSD metric. This result would indicate that the SSD metric is better able to capture controller workload than the DD metric, when tuning for a specific sector layout is not feasible.

6.1 Introduction

In previous chapters, the investigation of the Solution Space Diagram (SSD) metric in static and dynamic situations has shown that the metric correlates with workload at the same level or higher than commonly-used complexity metrics, such as the number of aircraft. To more thoroughly investigate the applicability and potential advantages of the SSD metric, it is crucial to compare it with a widely accepted complexity metric: Dynamic Density (DD). In this chapter, the number of aircraft and the DD metric are compared to the SSD metric in terms of their correlations to controller workload. Specific attention will be paid to the transferability of these metrics across different sector layouts and group of controllers.

An overview of the methodology to perform this study is shown in Figure 6.1.

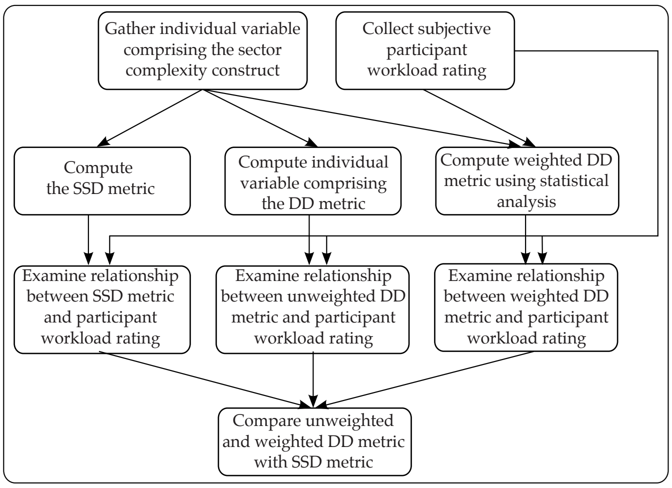


Figure 6.1: Work flow of SSD and DD metric comparison.

As the study described in this chapter relies on correlation analysis between the controller's workload ratings with the complexity metrics, the first step includes collecting subjective workload ratings throughout the experiment at regular time intervals to capture a workload profile for each controller. Secondly, based on the recorded aircraft parameters, such as such as position, speed, and heading, the SSD and unweighted DD metrics can be computed after a run. Linear regression analysis will then be performed to gather weighting coefficients corresponding to a number of Dynamic Variables (DV) to produce the weighted NASA DD metric that improves the correlations per individual. With all the information gathered, the comparison study between the number of aircraft and also both the unweighted and weighted NASA DD with the SSD can be facilitated.

6.2 Experimental Design

6.2.1 Subjects and Tasks

In the experiment, the participating eight male subjects with age between 29 and 51 (μ = 35.63, σ = 8.18), have all received an extensive Air Traffic Control (ATC) introductory course. As such, all subjects have a similar basic experience level in Air Traffic Management (ATM). The subjects were instructed to clear aircraft to their designated sector exit points and keep aircraft separated by at least 5 NM.

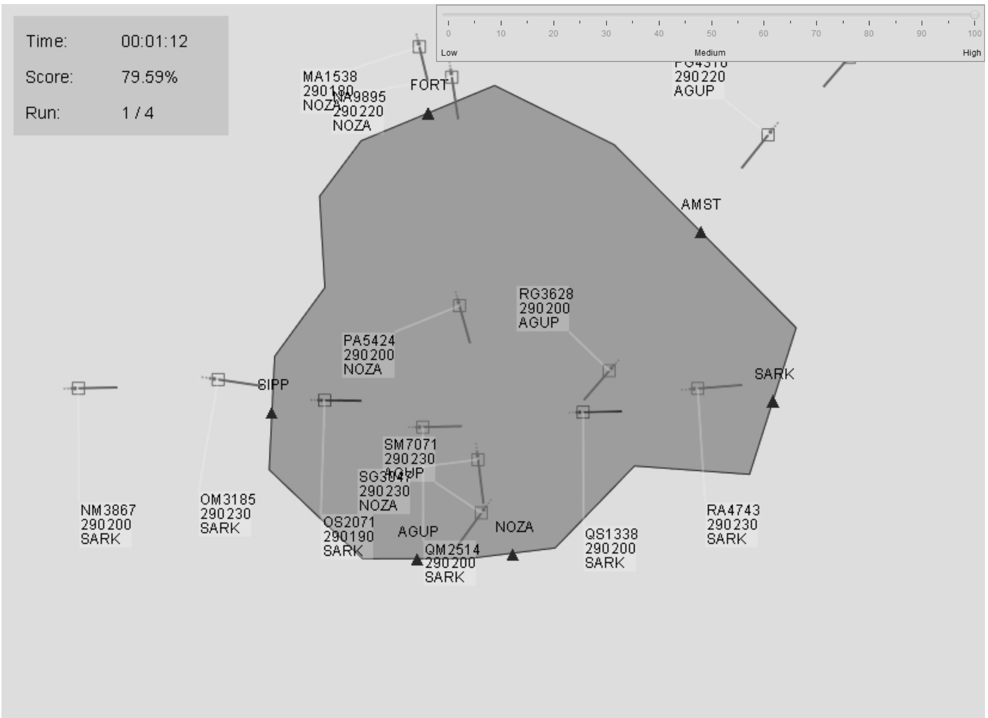


Figure 6.2: Experiment simulator.

All traffic was situated at FL290 and the function to change the altitude of aircraft was not enabled. Thus, the participants could only use heading and/or speed clearances to control aircraft. To support the controllers in their task, aircraft were color coded to indicate their course deviations and when they were in conflict. The unselected aircraft, which were headed towards their assigned exit point were colored in green, whereas unselected and deviating aircraft were colored in gray. Further, a selected aircraft was colored in white and would display an inner circle, indicating the 5 NM protected zone, and a green circle that indicated the current speed and a magenta circle and line, indicating the intended speed and heading clearance. In safely separating aircraft, a predicted loss of separation within 3 minutes (simulated-time) would trigger an aural alert and the involved aircraft in the conflict would be colored in red. Figure 6.2 shows an example of the simulator presented to the subjects.

Only aircraft which were inside the controlled sector could be given a speed and/or heading command. To control an aircraft, subjects first had to select an aircraft. Then, by dragging the heading line with the mouse to a new heading and/or scrolling the mouse scroll wheel up or down for speed change, the state of the aircraft could be changed. To confirm and implement a speed and/or heading change, the enter key had to be pressed.

During the experiment, the participants were asked to rate their perceived workload every 60 seconds. An automated stimulus provided a scale on the display that triggered the participants to rate their workload by means of clicking between 0 (low workload) and 100 (high workload). Unlike in Chapter 4 and 5, a mouse click on a scale that appeared on the same display (Figure 6.2) is presumed to provide subjects with a more direct and less intrusive workload rating measure than typing a number on a keyboard. The scale is also much finer grained, allowing the slightest change in workload to be captured.

6.2.2 Independent Variables

The independent variables in the experiment were the sector design, which had two levels (Sector 1 and Sector 2), and the incoming traffic sequence, which had four types (1 to 4).

Table 6.1: Sequence of scenario based on subject.

Subject	Group	Sector 1						Sector 2				
		Training		Incoming Traffic sequence				Training	Incoming Traffic sequence			
1	1	1	2	3	4	1	2	1	3	4	1	2
2	1	1	2	1	4	2	3	1	1	4	2	3
3	1	1	2	3	2	4	1	1	3	2	4	1
4	1	1	2	1	3	4	2	1	1	3	4	2
Subject	Group	Sector 2						Sector 1				
		Training		Incoming Traffic sequence				Training	Incoming Traffic sequence			
5	2	1	2	4	3	2	1	1	4	3	2	1
6	2	1	2	2	3	1	4	1	2	3	1	4
7	2	1	2	4	1	3	2	1	4	1	3	2
8	2	1	2	2	4	3	1	1	2	4	3	1

The independent variables provide a total number of (4 x 2 =) 8 experimental scenarios. The presentation order of the eight scenarios was randomized to counterbalance a possible order effect on the dependent measures in the experiment. Each incoming traffic sequence is different from another to prevent scenario recognition, such that recognition of a traffic pattern would not influence the dependent measures. Every subject had a scenario sequence assigned to him and this is shown in Table 6.1. Two groups of four controllers were designated, to investigate the effect of different group of controllers towards DD and SSD metrics. Both groups will have subjects of the same knowledge level regarding ATC.

6.2.3 Dependent Measures

The dependent measures in the experiment consisted of subject's control activity, number of aircraft properties, complexity measures, and subjective workload ratings.

i) Subject's control activity: The user control activity was measured by the number of speed commands (N_{speed}), the number of heading commands (N_{head}), the number of combined commands (N_{combi}) and the number of aircraft click (N_{air}).

ii) Number of aircraft properties: Two type of aircraft count were measured by the total number of aircraft within the currently controlled sector (N_{sum}) and the mean number of aircraft within the currently controlled sector (N_{mean}).

iii) Complexity measures: The complexity measures consisted of two DD metrics and the SSD area metric. Both DD metrics were measured every 60 seconds to match with the workload rating instances. The first DD metric, NASA DD Metric 1 ($NASA_1$) is based on research conducted by Chatterji & Sridhar (2001). For further details and calculation methods, readers are encouraged to refer to Chatterji & Sridhar (2001). The $NASA_1$ DD metric is calculated as follows:

$$DD = \sum W_i DV_i \quad (6.1)$$

The metric consisted of the following 16 Dynamic Variables (DV):

$NASA_1$ $DV1$: Number of aircraft
$NASA_1$ $DV2$: Fraction of climbing aircraft
$NASA_1$ $DV3$: Fraction of cruising aircraft
$NASA_1$ $DV4$: Fraction of descending aircraft
$NASA_1$ $DV5$: Horizontal proximity metric 1
$NASA_1$ $DV6$: Horizontal proximity measure 2
$NASA_1$ $DV7$: Horizontal proximity measure 3
$NASA_1$ $DV8$: Vertical proximity metric 1

$NASA_1$ $DV9$: Vertical proximity measure 2
$NASA_1$ $DV10$: Vertical proximity measure 3
$NASA_1$ $DV11$: Time-to-go to conflict measure 1
$NASA_1$ $DV12$: Time-to-go to conflict measure 2
$NASA_1$ $DV13$: Time-to-go to conflict measure 3
$NASA_1$ $DV14$: Variance of speed
$NASA_1$ $DV15$: Ratio of standard deviation of speed to average speed
$NASA_1$ $DV16$: Conflict resolution difficulty based on crossing angle

The second DD metric, NASA DD Metric 2 ($NASA_2$) calculation based on research by Laudeman et al. (1998) and Sridhar et al. (1998). For further details and calculation methods, readers are encouraged to refer to Laudeman et al. (1998) and Sridhar et al. (1998). For $NASA_2$, the weighted DD metric is gathered through its proposed equation as follows with TD as traffic density:

$$DD = \sum W_i DV_i + TD \tag{6.2}$$

The metric consisted of the following 8 DV, excluding traffic density:

$NASA_2$ $DV1$: Heading change
$NASA_2$ $DV2$: Speed change
$NASA_2$ $DV3$: Altitude change
$NASA_2$ $DV4$: Minimum distance 0-5 nm
$NASA_2$ $DV5$: Minimum distance 5-10 nm
$NASA_2$ $DV6$: Conflict predicted 0-25 nm
$NASA_2$ $DV7$: Conflict predicted 25-50 nm
$NASA_2$ $DV8$: Conflict predicted 40-70 nm

The original NASA DD metrics represented in researches by Chatterji & Sridhar (2001), Laudeman et al. (1998) and Sridhar et al. (1998) were constructed based on a 3-Dimensional (3D) airspace model. In gathering airspace and traffic factor to produce NASA DD metrics from a 2-Dimensional (2D) airspace model as used in this thesis, several DVs were canceled out from both $NASA_1$ and $NASA_2$ metric. These DVs are relevant to changes in altitude measures ($DV2$ and $DV4$ for $NASA_1$ metric and $DV3$ for $NASA_2$ metric) and also related to vertical proximities ($DV8$, $DV9$ and $DV9$ for $NASA_1$ metric).

The SSD area properties were calculated using the mean of SSD area (A_{mean}) of all aircraft within the sector (referred in this chapter as SSD). It is gathered using the following equation with A_{whole} representing the total area within minimum and maximum velocity-heading band of each individual aircraft within the sector and n being the number of aircraft within the sector. The SSD area properties were measured every 60 seconds to match with the workload rating instances.

$$A_{mean} = \frac{1}{n} \sum_{i=1}^{n} A_{whole_i}$$

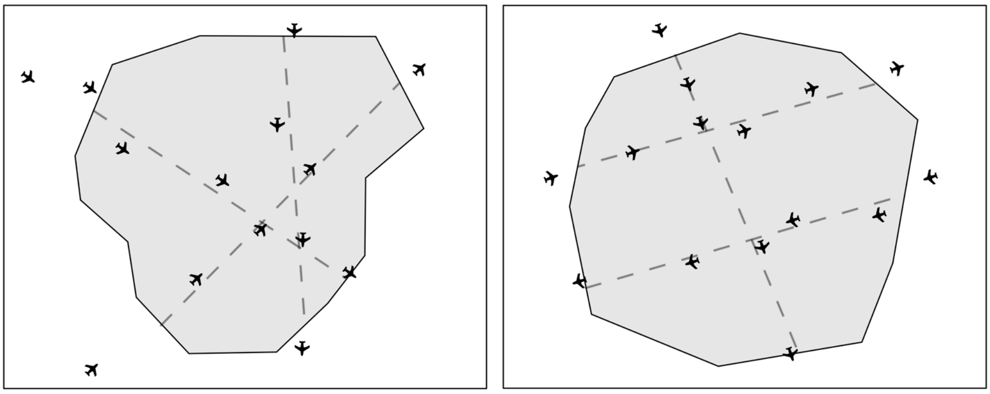
iv) Subject's workload ratings: The workload rating, measured on a zero to 100 scale, was provided by the subject every 60 seconds during the experiment run. In order to correct for inter-subject differences, Z-scores of the subjective ratings were used in the subsequent data exploration. This correction was performed by calculating the Z-scores for every test subject.

6.2.4 Sector Layout

The experiment scenarios were constructed based on the 'clearance to exit point task' with one type of aircraft on one flight level. There were three streams of incoming aircraft entering the sector. Apart from these three main similarities in sector designs, a number of differences were designed in both sectors in order to produce two different sectors. This is crucial in order to be able to test the metrics sensitivity to sector design. Figure 6.3 shows an example of the two sector designs used in this experiment.

The sector design variables can be observed from Figure 6.3 and the settings are detailed as follows:

1. Sector 1 has three **crossing points**, while Sector 2 has two crossing points.
2. Sector 1 has mixed combinations of the **intercept angle of traffic routes** of approximately 45°, 90° and 120°. Whereas Sector 2 has two approximately 90° crossing angles.
3. The two sectors had a different pattern in **crossing point clusters**. That is, Sector 1 had more clustered crossing points near the sector border, whereas Sector 2 had a less clustered intersection points with the two crossing points having ample spacing between them.
4. Both sector also had a different pattern in the **clustering of entry and exit points**. Sector 1 has all exit points on the right hand side of the sector, whereas Sector 2 has entry and exit points at both sides of the sector.
5. **Different sector shape** were designed for both sectors. Sector 1 has a more odd polygon shape, whereas Sector 2 has a more regular polygon shape.
6. The two sectors had different **sector area** properties. Sector 1 has an area of approximately 30% less than Sector 2. Sector 1 has a total area of $7000nm^2$, whereas Sector 2 has a total area of $10400nm^2$.



(a) Sector 1 **(b)** Sector 2

Figure 6.3: Sector design and the traffic flow assignment.

6.2.5 Procedure

Before commencing the experiment, subjects were briefed on the nature of the experiment, the goals to be achieved and the simulator that was used for the experiment. Each participant had to complete two blocks of scenarios. In each block, four scenarios were required to be controlled and each scenario lasted for 25 minutes. Each block is preceded with a training scenario that lasted for 10 minutes. During the course of the experiment, participants were asked to indicate their workload using a scale that appeared on top of the simulator screen. The experiment was run at 4 times real-time, similar to previous experiment procedures.

6.2.6 Hypotheses

The experiment was intended to compare the SSD metric and other sector complexity measures, namely the number of aircraft and DD. The hypotheses concerning the SSD capacity as a reliable and objective workload measure were as follows:

1. The SSD results in a higher correlation with controller workload than the number of aircraft and the unweighted NASA DD metric.
2. The SSD has comparable workload correlations to the weighted DD metric.
3. The SSD correlations are less sensitive to sector and controller changes than the weighted DD metric.

6.3 Results

The overall data consisted of data gathered within the 25 minutes experiment duration. However, only data after the first 3 minutes up until 23 minutes (resulting in a window of 20 minutes) were used in the analysis as illustrates in Figure 6.4. This was done to rule out 'fade in' and 'fade out' effects in terms of workload and alertness of the controllers. In this result section, effects were considered significant at a probability level $p \leq 0.05$, where p is the probability that the null hypothesis is true.

Figure 6.4: Period where data is gathered in the experiment.

The section will start with the discussion on the effect of assigning multiple incoming traffic sequence to represent single sector. Then, the effects of the two sectors towards controller's behavior and sector complexity measure are laid out. Next, in order to do a comparison between complexity measures and workload rating, correlation analysis of different sector complexity measure (number of aircraft, unweighted NASA DD and SSD) towards subjective workload rating is carried out. Then, a regression analysis was performed on both NASA DD metrics. Correlations between weighted DD metric and subjective workload rating was performed in order to again compare different complexity measures. To demonstrate DD sensitivity towards different sector designs and groups of controllers, the weight coefficients gathered from linear regression analysis was swapped between sectors and also between groups of controllers.

Different incoming traffic sequence in one sector

In the design of the experiment, a total of four different incoming traffic sequences were assigned to each sector. By using four repetitions, more traffic data could be gathered. For each sector, the initial traffic design have approximately the same total number of aircraft with varying incoming aircraft sequence. With different controllers control strategy, the resulting number of aircraft within the sector at a certain point in time can also change. Figure 6.5 showed the resulting distribution of mean of number of aircraft for each rating sequence instance based on different incoming traffic sequence and sector (for all subjects).

It is observed from Figure 6.5 that the fluctuations of the number of aircraft were almost identical for different traffic sequences within a particular sector. Based on initial Friedman analysis (test statistic χ^2) on different scenarios for Sector 1 and 2, it is gathered that both sectors did not show any significant difference in the number of aircraft present for different incoming traffic sequence with χ^2 = 0.750, p = 0.861 for Sector 1 and χ^2 = 0.494, p = 0.920 for Sector 2.

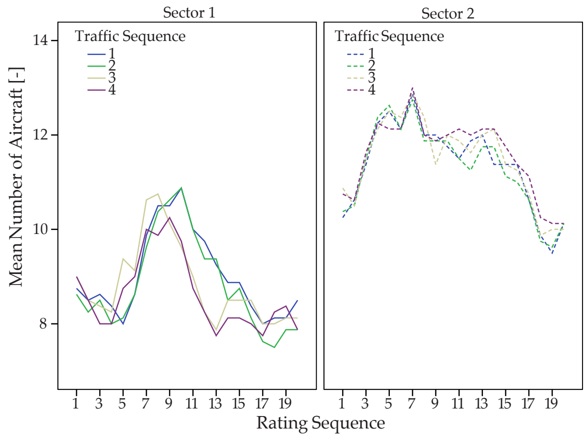


Figure 6.5: Mean number of aircraft per workload rating sequence.

However, before making the assumption that the data from the four traffic sequences within each sector can be jointly analyzed on a sector level, an analysis on the controller's workload rating in different traffic scenarios was conducted. To investigate this, statistical analysis using Friedman's non-parametric ANOVA and post-hoc analysis using the Wilcoxon-Signed Rank test (test statistic Z) were chosen.

Based on initial Friedman analysis on different scenarios for Sector 1 and 2, it is gathered that the controller's workload rating produced no significant difference for different incoming traffic sequence with χ^2 = 0.300, p = 0.960 and χ^2 = 1.050, p = 0.789 in Sector 1 and Sector 2, respectively.

Thus, based on these findings, we can assume that all four traffic to be similar to one another and differences in the traffic are not of an influence to the level of workload experience by the controller. For that reason, from this point forward, analysis were made directly based on different sector with the assumption that different incoming traffic sequence is considered to be one single group of data.

Different sector assignment

To enable cross-sector analysis, it is important that both sectors represent different levels of complexity for controllers. This section explores the control behavior (N_{speed}, N_{head}, N_{combi} and N_{air}), complexity measures ($NASA_1$, $NASA_2$, SSD, N_{sum} and N_{mean}) and workload ratings between different sectors.

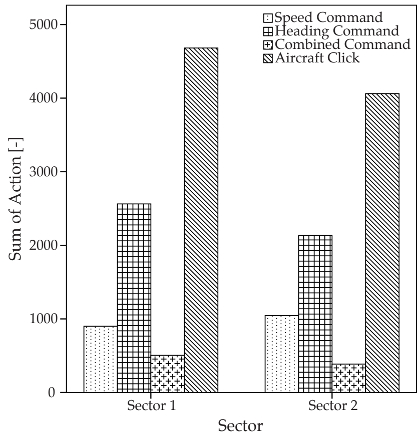
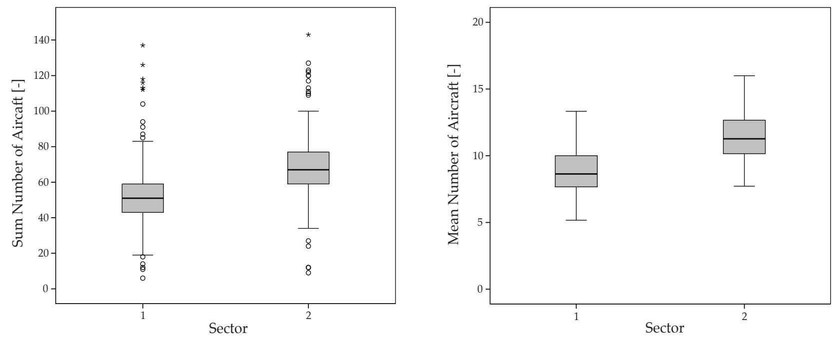


Figure 6.6: Sum of command based on different sectors.

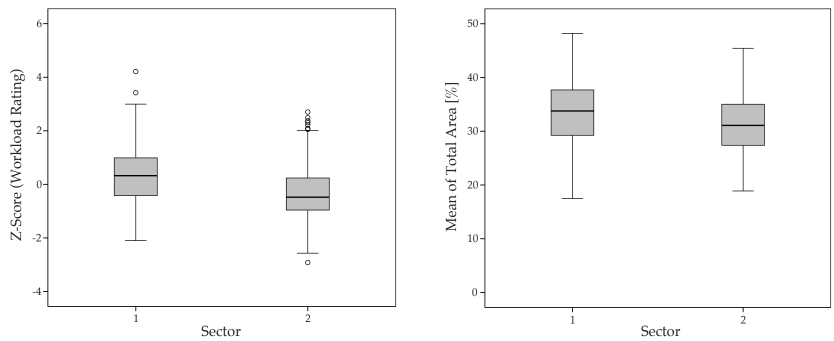
Based on the analysis, it is found that a significant effect ($p < 0.05$) of different sector designs can be observed for certain subject's control behavior, especially in regards to heading commands and aircraft clicks. Figure 6.6 illustrate differences in command patterns, which indicate different strategies used for both sectors. Sector 1, having more crossings than Sector 2, showed a higher N_{head}. On the other hand, a higher N_{speed} was found in Sector 2, in response to the two parallel route structure. The number of aircraft clicks is also higher in Sector 1 compared to Sector 2 as a result of a more difficult aircraft maneuverings in responds to multiple crossings in cross proximity, which can be observed from Figure 6.3a.

The two sectors also showed a significant difference in both number of aircraft (p = 0.012 for both N_{sum} and N_{mean}) and workload ratings (p = 0.025). A higher number of aircraft was observed in Sector 2 (Figure 6.7a and 6.7b). However, for the controller workload ratings, the opposite was found (Figure 6.8a), where Sector 2 was observed to produce a lower workload rating. This shows that the number of aircraft within a sector is not the only factor that induces higher workload ratings. In this experiment, differences in sector design have influenced controller to rate Sector 1 with a higher workload rating even when the number of aircraft is lower. This is also consistent with the concept of having a different maximum number of aircraft per sector.



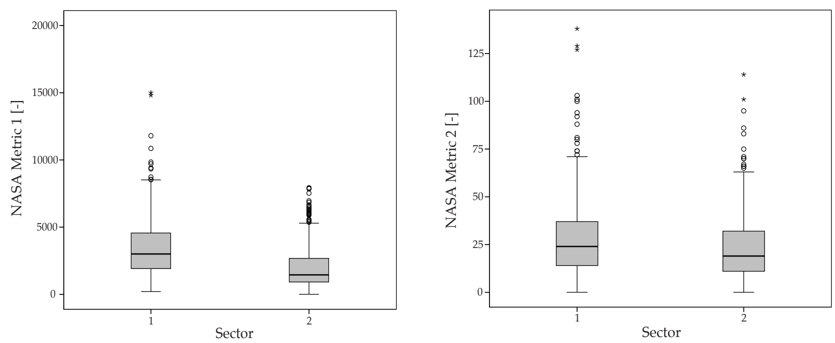
(a) N_{sum} **(b)** N_{mean}

Figure 6.7: Number of aircraft based on different sectors.



(a) Workload rating **(b)** SSD

Figure 6.8: Workload rating and SSD area properties based on different sectors.



(a) $NASA_1$ **(b)** $NASA_2$

Figure 6.9: NASA DD metrics based on different sectors.

Figure 6.8b and 6.9 show the SSD and DD behavior towards different sector designs, respectively. Based on the figures, both sector complexity measures showed a similar pattern with Sector 2 consistently showing lower values than Sector 1 (p = 0.012 for both SSD and DD metrics), corresponding to the workload rating results presented earlier. It is concluded that both sectors indeed represent different levels of sector complexity. Thus, based on that fact, it is deemed possible to investigate the sensitivity of sector complexity measures towards different sector designs.

6.3.1 Unweighted Correlation Analysis

In current practice, air traffic complexity is generally based on the number of aircraft (Hilburn, 2004, Sridhar et al., 1998). However, to investigate whether either the number of aircraft or the NASA or the SSD metric would best represent controller workload, a correlation analysis between the number of aircraft, the unweighted NASA DD and the SSD metrics with respect to subjective workload ratings were performed. The analysis was conducted using the Kendall's tau correlation analysis (test statistic R) based on data gathered during a 20 minutes experiment run.

Sector-based analysis

The analysis of unweighted NASA DD metrics was made based on the assumption that all DV weighting coefficient are equal and were all assigned as 1. The unweighted NASA DD metrics in this section were calculated using Equation (6.1) and (6.2).

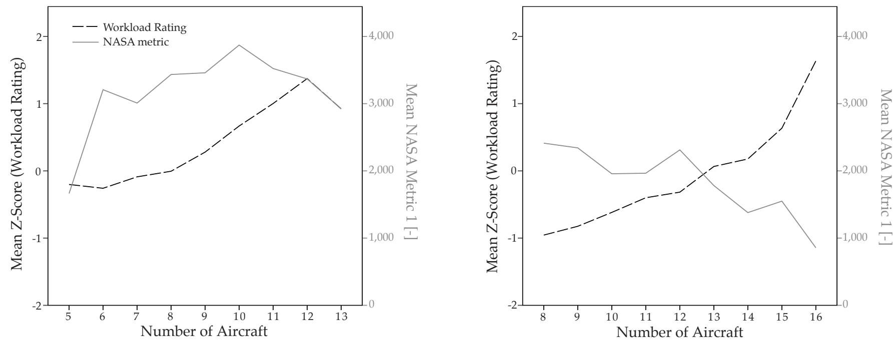
Table 6.2: Correlation coefficient between workload rating and sector complexity measures based on different sector.

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

Table 6.2 showed results of the correlation analysis between workload rating and sector complexity measures. Based on the results, SSD showed the highest correlation with workload rating (highest correlation in **bold**). N_{mean} is second in line as a good sector complexity measure which demonstrates that indeed the number of aircraft

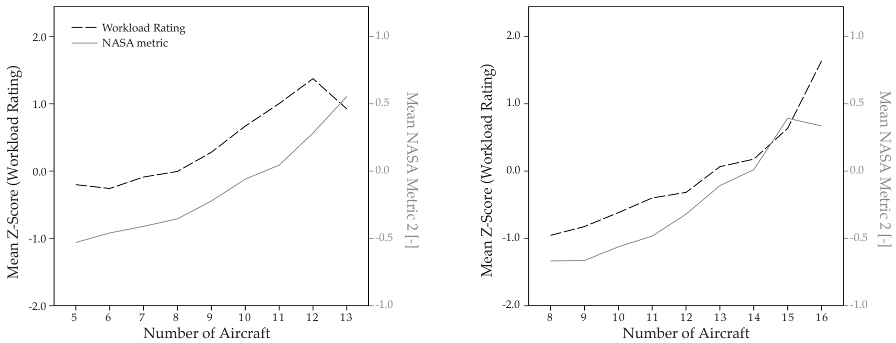
is one of the most important sector complexity variable that influences controller's workload.

Figure 6.10, 6.11 and 6.12 showed plots of unweighted NASA DD and SSD metric compared to workload rating based on number of aircraft, respectively. The plots were intended to illustrate how workload rating behave towards number of aircraft and also how unweighted NASA DD and SSD metric behave in responds to the same number of aircraft.



(a) Sector 1 **(b)** Sector 2

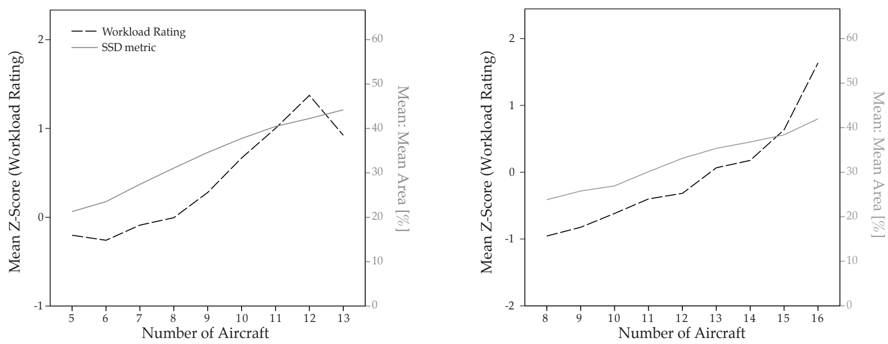
Figure 6.10: Unweighted $NASA_1$ based on different sectors.



(a) Sector 1 **(b)** Sector 2

Figure 6.11: Unweighted $NASA_2$ based on different sectors.

Based on Figure 6.10a and 6.10b, $NASA_1$ plots did not show a pattern that is closely related to workload rating. Other sector complexity measures such as $NASA_2$ (Figure 6.11) and SSD (Figure 6.12) showed a more resembling pattern that of the workload rating.



(a) Sector 1 **(b)** Sector 2

Figure 6.12: SSD area properties based on different sectors.

Controller group analysis

The same result was observed when looking at different group of controllers with SSD showing highest correlation with workload rating (highest correlation in **bold**). Table 6.3 shows the results of unweighted NASA DD, number of aircraft and SSD metric.

Table 6.3: Correlation coefficient between workload rating and sector complexity measures based on different group.

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

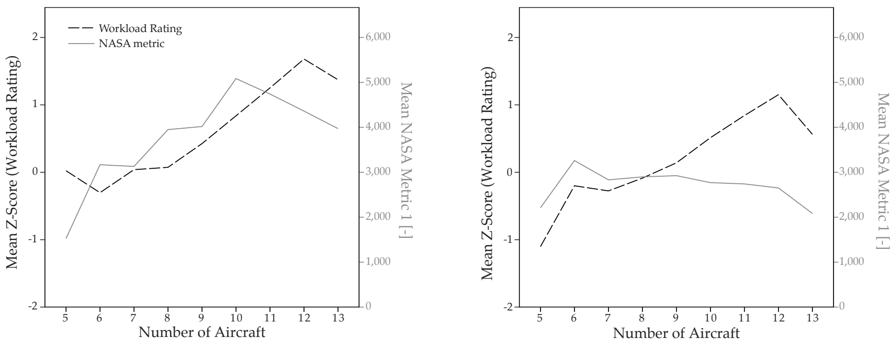
$NASA_2$	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	0.269	0.261	0.306	0.204
p	<0.001	<0.001	<0.001	<0.001

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

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

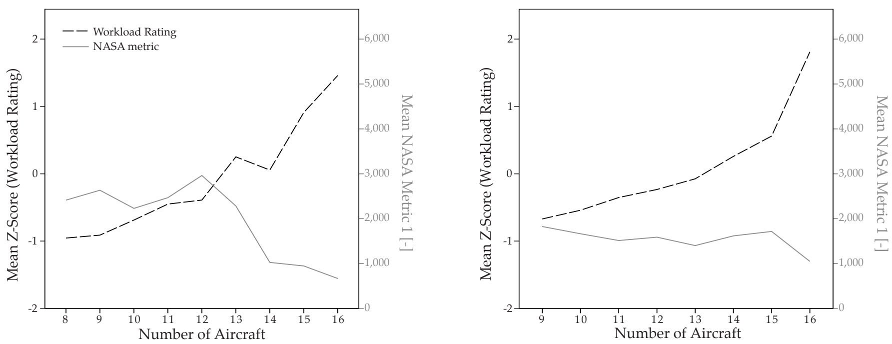
SSD	Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2
R	**0.362**	**0.335**	**0.341**	**0.232**
p	**<0.001**	**<0.001**	**<0.001**	**<0.001**

To investigate the effect of before and after regression analysis on NASA DD metric, Figure 6.13 to 6.16 were illustrated to demonstrate the behavior of unweighted NASA DD metrics in comparison with workload rating. Figure 6.17 and 6.18 were illustrated to demonstrate the behavior of SSD metrics in comparison with workload rating.



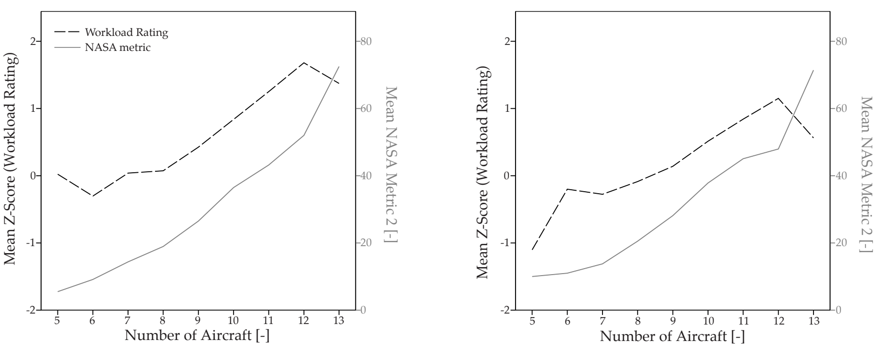
(a) Group 1 **(b)** Group 2

Figure 6.13: Unweighted $NASA_1$ based on different group of controllers (Sector 1).



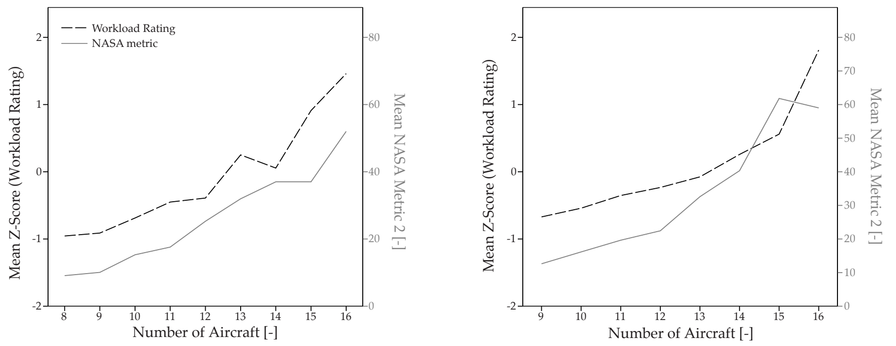
(a) Group 1 **(b)** Group 2

Figure 6.14: Unweighted $NASA_1$ based on different group of controllers (Sector 2).



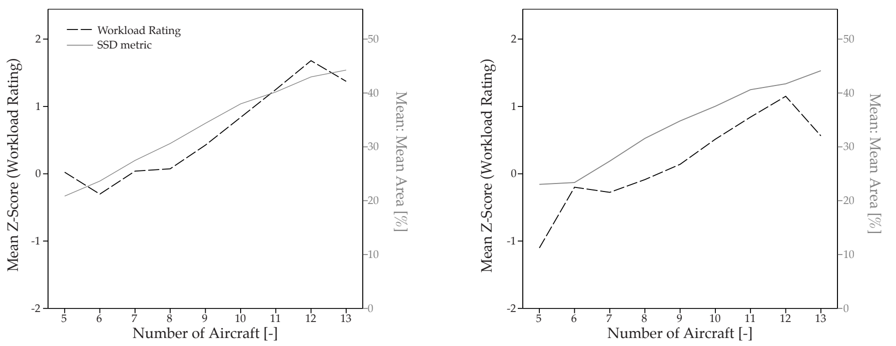
(a) Group 1 **(b)** Group 2

Figure 6.15: Unweighted $NASA_2$ based on different group of controllers (Sector 1).



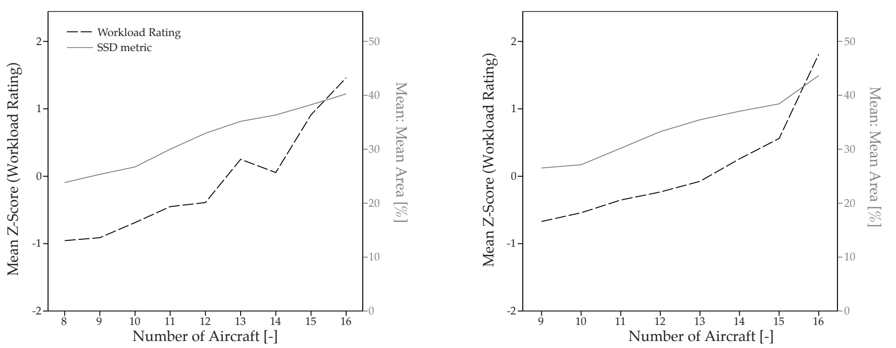
(a) Group 1 **(b)** Group 2

Figure 6.16: Unweighted $NASA_2$ based on different group of controllers (Sector 2).



(a) Group 1 **(b)** Group 2

Figure 6.17: SSD area properties based on different group of controllers (Sector 1).



(a) Group 1 **(b)** Group 2

Figure 6.18: SSD area properties based on different group of controllers (Sector 2).

Comparing Figure 6.13 and 6.14 to Figure 6.15 and 6.16, $NASA_2$ showed a more resembling pattern that of the workload rating, which is also shown through the correlation analysis results in Table 6.3. As for the SSD metric, a more resembling pattern that of the workload rating can be observed from Figure 6.17 and 6.18.

6.3.2 Weighted Correlation Analysis

In this section, the unweighted NASA DD metric were fixed to the workload rating data, using the linear regression method, resulting in a fitted weighted NASA DD metric. In principle, the weighted NASA DD metric should correlate better than the unweighted ones. The regression analysis was conducted based on different sector and also based on different group of controllers. This is done in order to investigate whether the weighted NASA DD metric is consistently better than the SSD regardless of different sector or group of controller. Also, in the subsequent section, analysis on the transferability of the weighted NASA DD metric across different sector design or group of controllers will be carried out.

Sector-based analysis

First, linear regression analysis were conducted on the basis of different sectors. Based on the analysis, a number of significant variables were identified. Variables that computed regression weights were small and non significant were removed from the equation that was used to compute the end DD. The weighted NASA DD metric were constructed based on the coefficient individual contribution (b-value), representing the weighting factor for each DV. By replacing the significant b-value into equation (6.1) and (6.2), the NASA DD model can be defined as follows with the corresponding DV detailed in subsection 6.2.3:

1) Sector 1:

$$\begin{aligned} NASA_1 &= 1.134 + 0.191 * DV1 - 0.738 * DV3 + 7.301 * DV6 \\ &\quad + 0.534 * DV11 + 0.0003 * DV14 - 1.819 * DV15 \\ &\quad - 0.0003 * DV16 \\ NASA_2 &= -0.466 + 0.111 * DV1 + 0.111 * DV2 + 0.023 * DV5 + TD \end{aligned}$$

2) Sector 2:

$$\begin{aligned} NASA_1 &= -0.761 * DV3 + 9.902 * DV5 + 3.043 * DV6 + 1.750 * DV7 \\ NASA_2 &= -0.844 + 0.098 * DV1 + 0.036 * DV5 + 0.012 * DV6 + TD \end{aligned}$$

For both sectors, the $NASA_1$ DD metric are defined as having different significant DV, which are included in the end DD equation. In Sector 1, the significant DV are focused more to the variables related to aircraft horizontal proximity ($DV6$), speed ($DV14$ and $DV15$) and intercept angle ($DV16$), whereas in Sector 2, only variable concerning horizontal proximity ($DV5$ to $DV7$) are found to be significant. It is also concluded that the number of aircraft has shown a significant effect for Sector 1, but not in Sector 2.

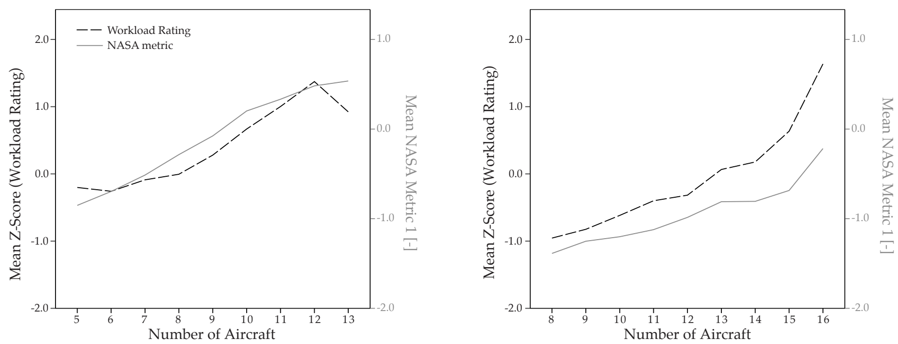
For the $NASA_2$ DD metric, the speed change variable ($DV2$) showed to be significant in Sector 1, but not in Sector 2. However, in both sectors, variable concerning heading change ($DV1$) and horizontal proximity ($DV5$) were found to be significant. Differences in variables that influences the NASA DD model for both sector showed that different sector design demand for different weighted NASA DD metric.

The correlation between the resulting weighted DD and workload rating were gathered again using Kendall's tau correlation coefficient. Results are gathered and presented in Table 6.4. For SSD correlation data refer to Table 6.2. Based on the result, $NASA_1$ for Sector 1 and $NASA_2$ for Sector 2 have higher correlation than SSD (highest correlation in **bold**). It is observed that weighted $NASA_1$ showed an increases in correlation on both sector if compared to unweighted $NASA_1$. However weighted $NASA_2$ showed a lower correlation in Sector 1 and a higher correlation in Sector 2 compared to unweighted $NASA_2$.

Table 6.4: Correlation coefficient between workload rating and weighted NASA DD metric (sector-based analysis).

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

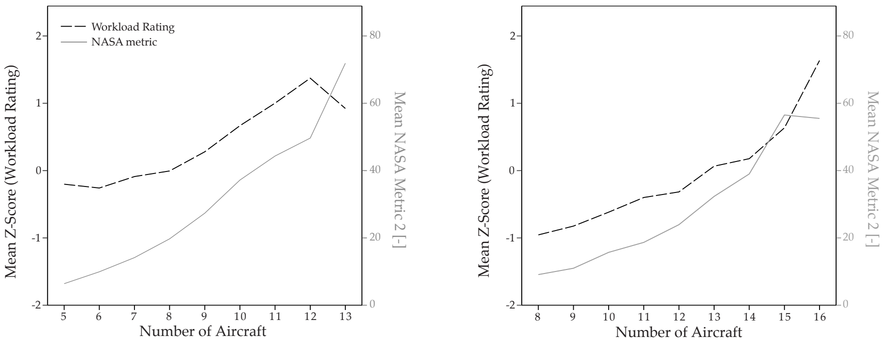
*correlation at a lower level than unweighted NASA DD metric



(a) Sector 1 **(b)** Sector 2

Figure 6.19: Weighted $NASA_1$ based on different sectors.

Figure 6.19 and 6.20 showed weighted NASA DD with workload rating also against the number of aircraft. This can be compared with the initial unweighted NASA DD from Figure 6.10 and 6.11 where the plots of weighted $NASA_1$ and $NASA_2$ have improved to a plot that better matches the workload rating in Figure 6.19 and 6.20.



(a) Sector 1 **(b)** Sector 2

Figure 6.20: Weighted $NASA_2$ based on different sectors.

Controller group analysis

Secondly, analysis based on different group of controllers within the same sector was conducted to investigate how different a group of people with the same level of knowledge on ATM can behave. Linear regression analysis was conducted and only significant variables were used to compute the end DD. By replacing the significant b-value into equation (6.1) and (6.2), we can define the NASA DD model as follows:

1) Sector 1 and group 1:

$$NASA_1 = 1.833 + 0.239 * DV1 - 0.841 * DV3 + 7.319 * DV6 + 0.753 * DV11 + 0.0002 * DV14 - 0.0004 * DV16$$
$$NASA_2 = -0.487 + 0.106 * DV1 + 0.116 * DV2 + 0.025 * DV5 + TD$$

2) Sector 1 and group 2:

$$NASA_1 = 0.654 - 0.685 * DV3 + 6.742 * DV6 + 0.0003 * DV14 - 2.223 * DV15$$
$$NASA_2 = -0.433 + 0.124 * DV1 + 0.097 * DV2 + 0.0.022 * DV5 + TD$$

3) Sector 2 and group 1:

$$NASA_1 = -0.702 * DV3 + 2.463 * DV7$$
$$NASA_2 = -0.899 + 0.100 * DV1 + 0.051 * DV5 + 0.030 * DV6 + TD$$

4) Sector 2 and group 2:

$$NASA_1 = 0.270 * DV1 - 0.808 * DV3$$
$$NASA_2 = -0.844 + 0.142 * DV1 + 0.134 * DV2 + 0.019 * DV5 - 0.024 * DV7 + TD$$

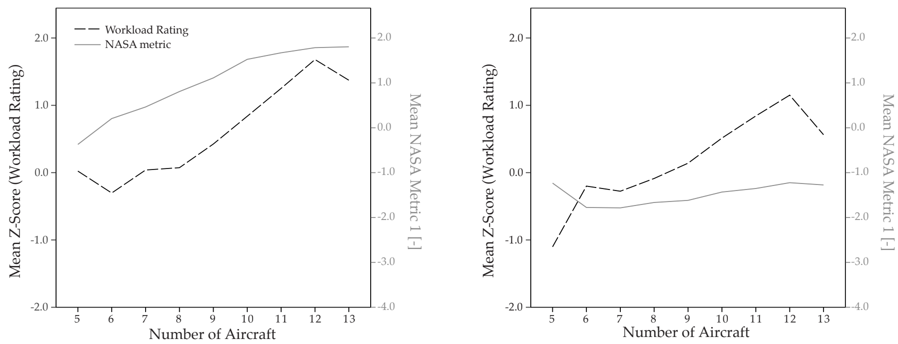
For the $NASA_1$ DD metric, in Sector 1, both groups showed a significant effect towards DVs which are related to the aircraft horizontal proximity ($DV5$ to $DV7$) and speed ($DV14$ and $DV15$). However, DVs which are related to the time-to-go to conflict ($DV11$ to $DV13$) and aircraft intercept angle ($DV16$) also played a role in the end DD metric for group 1, but not for group 2. In Sector 2, significant effect were found for aircraft horizontal proximity variables ($DV5$ to $DV7$), but only for group 1.

For the $NASA_2$ metric, in Sector 1, both groups display the same behavior, but in Sector 2, the speed change variable ($DV2$) is a significant factor in determining the end DD, but only for group 2. Differences in variables that influence the NASA DD model for both groups, showed that different group of controllers, demand for different weighted NASA DD metric as a result of differences in controllers behavior towards a particular sector. Workload addresses the subjective demand experienced by the operator in the performance of a task. It is influenced by operator-centered factors like skill, strategy, and experience. In the experiment sequence, group 1 has initially started with Sector 1, followed by Sector 2. Whereas group 2 has experienced the opposite situation. Thus, the difference in level of experience will effect the controller's strategy. This can be seen through the weighted NASA DD metric.

Table 6.5: Correlation coefficient between workload rating and weighted NASA DD metric (controller group-based analysis).

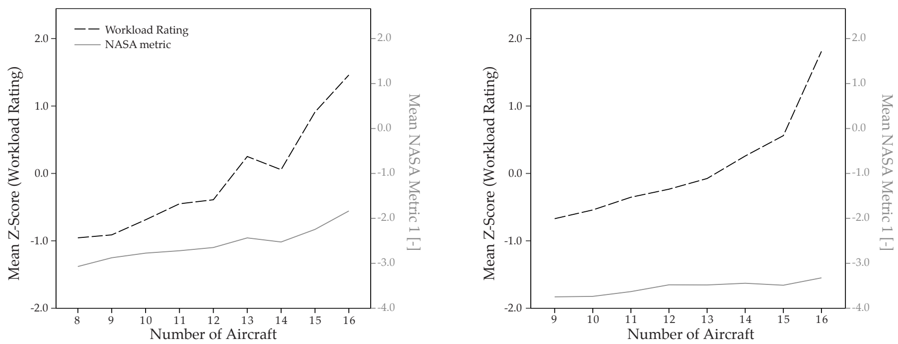
	$NASA_1$				$NASA_2$			
	Sector 1		Sector 2		Sector 1		Sector 2	
	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
R	**0.429**	0.229	0.217	0.111	0.358	**0.352**	**0.348**	**0.312**
p	<**0.001**	<0.001	<0.001	0.003	<0.001	<**0.001**	<**0.001**	<**0.001**

The correlation between the end weighted NASA DD and workload rating were re-analyzed using Kendall's tau correlation coefficient and the results are presented in Table 6.5. Based on the results, only $NASA_1$ in Group 1 for Sector 1 has higher correlation with workload rating compared to the SSD metric. However, for $NASA_2$, only Group 1 for Sector 1 has lower correlation with workload rating compared to the SSD metric, whereas other groups have higher correlation. For comparison, Table 6.3 presented correlation analysis of SSD with workload rating based on different group of controllers. It is also observed that both weighted $NASA_1$ and $NASA_2$ have better correlation than the respective unweighted NASA DD metrics.



(a) Group 1 **(b)** Group 2

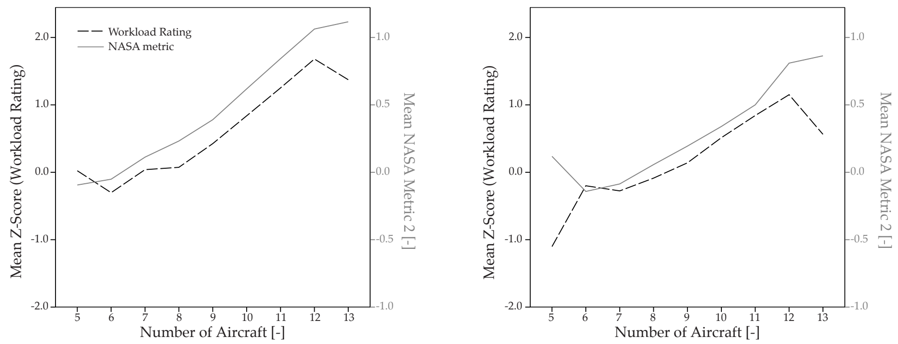
Figure 6.21: Weighted $NASA_1$ based on different group of controllers (Sector 1).



(a) Group 1 **(b)** Group 2

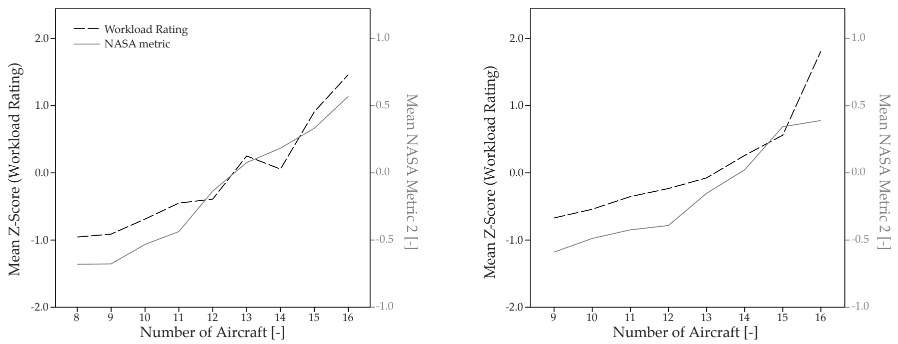
Figure 6.22: Weighted $NASA_1$ based on different group of controllers (Sector 2).

Figure 6.21 to 6.24 illustrate the relation between workload rating and weighted NASA DD metric as sector complexity measures. This can be compared with the initial unweighted NASA DD from Figure 6.13 to 6.16 where the plots of weighted $NASA_1$ and $NASA_2$ have improved to a plot that matches better the workload rating in Figure 6.21 to 6.24.



(a) Group 1 **(b)** Group 2

Figure 6.23: Weighted $NASA_2$ based on different group of controllers (Sector 1).



(a) Group 1 **(b)** Group 2

Figure 6.24: Weighted $NASA_2$ based on different group of controllers (Sector 2).

6.3.3 Transferability Analysis

In addition to the weighted NASA DD analysis, to demonstrate that the weighting coefficient only serves a certain sector or group of controllers, a cross analysis of NASA DD metric between different sector and controller group were carried out.

Cross-sector transferability

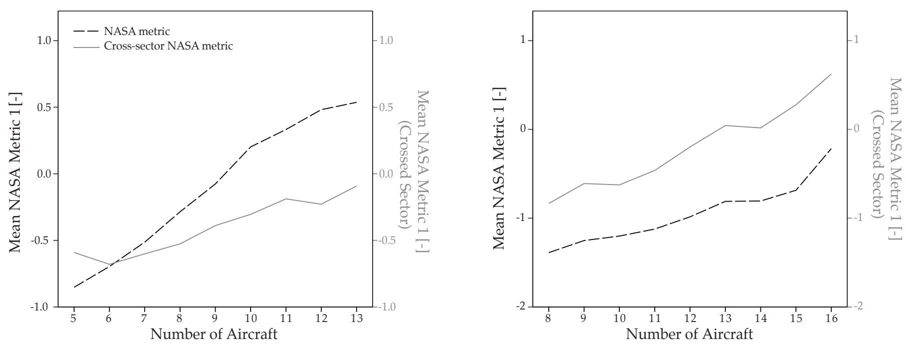
Firstly, cross-sector analysis was conducted by applying the weighting coefficient gathered in Sector 1 to Sector 2 and vice versa. Based on result in Table 6.6, only $NASA_2$ for Sector 1 showed a higher correlation level than the original correlation value. Others showed lower correlation level. However, both $NASA_1$ and $NASA_2$ showed lower correlation than SSD metric sector complexity measure.

Table 6.6: Correlation coefficient between workload rating and cross-sector weighted NASA DD metric.

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

*correlation at a higher level than weighted NASA DD metric

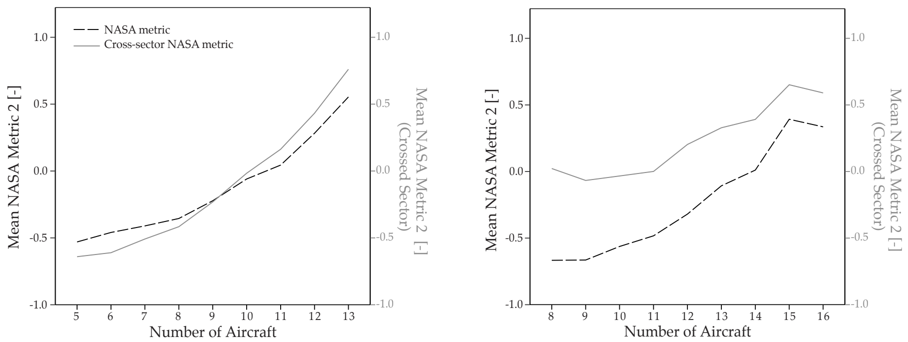
As observed in Table 6.6, the $NASA_2$ DD metric for Sector 1 showed a higher correlation level than the original correlation value that can be observed in Table 6.4. However, it should also be made aware that for Sector 1, the weighted $NASA_2$ DD metric (Table 6.4) showed a lower correlation compared to the unweighted $NASA_2$ (Table 6.2). As speed change variable ($DV2$) was present in $NASA_2$ DD metric for Sector 1 but not in Sector 2, outliers within the variable might have changed the output that the linear regression analysis produces and reduces the predictive accuracy of the weighted $NASA_2$ DD metric. However, the speed change variable was not removed from the weighted $NASA_2$ DD metric equation as it can also represent a result of different control strategy for different sector designs.



(a) Sector 1 **(b)** Sector 2

Figure 6.25: Difference of weighted $NASA_1$ and cross-sector weighted $NASA_1$.

To illustrate how different weighting coefficients influence the weighted NASA DD value, plots of original NASA DD metric value towards number of aircraft were shown together with the cross-sector NASA DD metric value at the same scale. This is illustrated in Figure 6.25 and 6.26. The fact that differences between original and cross-sector value are evident shows that regression analysis needs to be done for each corresponding sector before the NASA DD metric can be used as a sector complexity measure efficiently.



(a) Sector 1 **(b)** Sector 2

Figure 6.26: Difference of weighted $NASA_2$ and cross-sector weighted $NASA_2$.

To assess SSD metric sensitivity towards different sector and compare its behavior with NASA DD metric's, a scatter plot of workload rating towards sector complexity measure were illustrated in Figure 6.27. Based on Figure 6.27 the distribution of data based on all subjects for both Sector 1 and Sector 2 in SSD is almost identical. It is also observed that Sector 1 has a higher workload rating. However, this is accompanied by higher SSD area properties. Thus, showing that differences between sectors do not result in differences in how the SSD behaves towards workload rating.

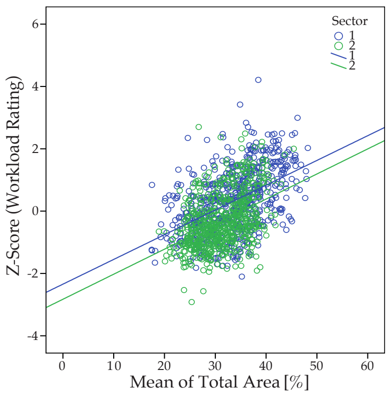


Figure 6.27: Scatter plot of workload rating and SSD area properties based on different sectors.

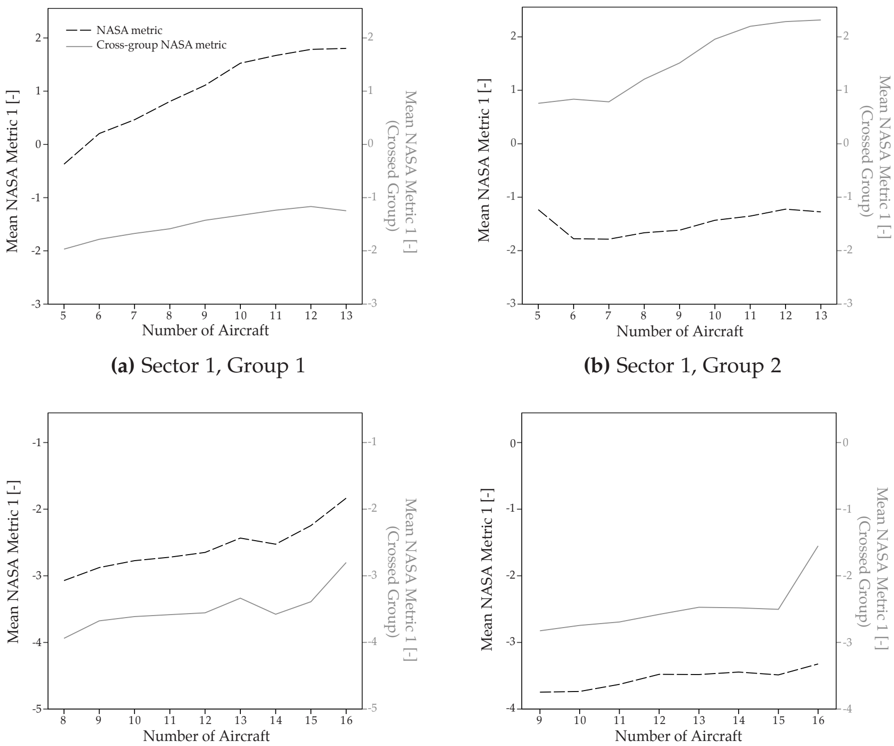
Cross-group transferability

Secondly, the effect of using weighting coefficient of different group of controllers towards another group of controllers were investigated. Correlation analysis were conducted and based on the result in Table 6.7, several cross-group weighted NASA DD value have correlation at a higher level than SSD and are highlighted in **bold**. Cross-group weighted NASA DD which has correlation at a higher level than its original weighted NASA DD value are highlighted with a '*'.

Table 6.7: Correlation coefficient between workload rating and cross-group weighted NASA DD metric.

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

*correlation at a higher level than weighted NASA DD metric



(a) Sector 1, Group 1

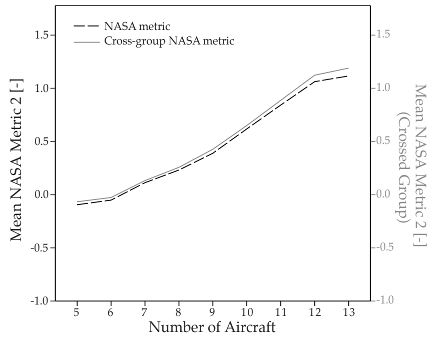
(b) Sector 1, Group 2

(c) Sector 2, Group 1

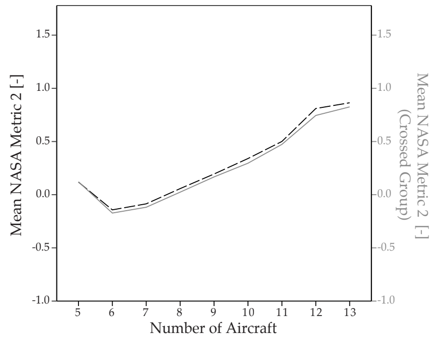
(d) Sector 2, Group 2

Figure 6.28: Difference of weighted $NASA_1$ and cross-group weighted $NASA_1$ based on different group of controllers and sectors.

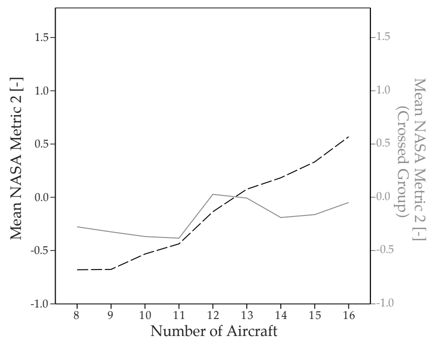
It can be observed that a mixed of higher and lower levels of correlation than the weighted $NASA_1$ DD metric were gathered for the $NASA_1$ cross-group analysis based on different groups of controllers. This is similar to the findings in cross-sector transferability analysis of $NASA_2$ DD metric. It is believed that outliers within the number of aircraft variable ($DV1$), which plays a role in the regression equation for Group 2, might have reduced the level of correlation between weighted $NASA_1$ DD metric and workload ratings. Thus, in cross-group analysis, when the variable is no longer included within the $NASA_1$ DD metric equation, improvement in the correlation coefficient can be observed. However, the number of aircraft variable was not removed from the initial weighted $NASA_1$ DD metric equation as it can also represent a result of different control strategy for different groups of controllers.



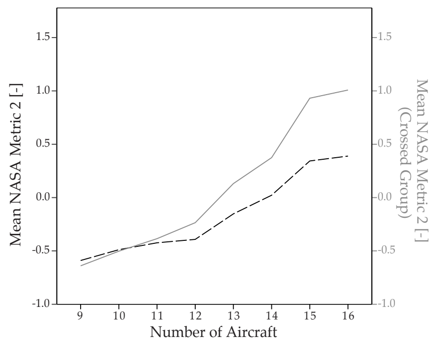
(a) Sector 1, Group 1



(b) Sector 1, Group 2



(c) Sector 2, Group 1



(d) Sector 2, Group 2

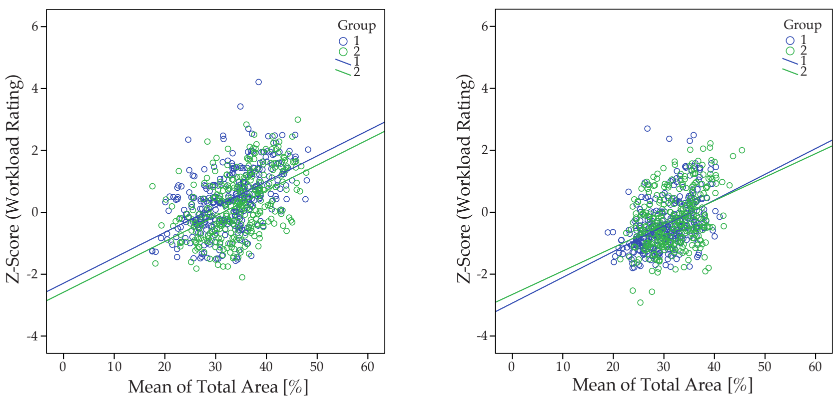
Figure 6.29: Difference of weighted $NASA_2$ and cross-group weighted $NASA_2$ based on different group of controllers and sectors.

This has indicated that the NASA DD metric is sensitive not only to differences in sector design but also sensitive towards different groups of assigned controllers.

Figure 6.28 and 6.29 showed how both $NASA_1$ and $NASA_2$ can be different when the weighting value of different group of controllers were used on a group of controllers.

It is also observed that $NASA_1$ showed a bigger difference in Sector 1 when the weighting factor of one group is transferred to another. This can be traced back to the end NASA DD equation in previous section with Sector 1 having more DV which are considered significant for group 1. For $NASA_2$, the opposite occurs with Sector 2 showed to have a bigger difference when the equation is transferred. The same rationale present for $NASA_2$.

The cross-group analysis reveals that overall, both NASA DD metric is sensitive towards different group of controllers. The effect of different group is more apparent in Sector 1 for $NASA_1$ DD metric, and in Sector 2 for $NASA_2$ DD metric. However, it is also observed that for Sector 1, the $NASA_2$ metric showed to be less sensitive towards different groups of controllers than Sector 2. This has shown that both NASA DD metric response differently to differences in sector design.



(a) Sector 1 **(b)** Sector 2

Figure 6.30: Scatter plot of workload rating and SSD area properties based on different group of controllers and sectors.

Figure 6.30 illustrates the relation between workload rating and the SSD metric as sector complexity measures. Based on the plots, it is observed that SSD metric showed little differences on the distribution of SSD data between groups for both sectors.

6.4 Discussion

This chapter compares the proposed metric, SSD with known metrics such as the number of aircraft and NASA DD metric gathered from research by Laudeman et al. (1998), Sridhar et al. (1998) and Chatterji & Sridhar (2001). Multiple scenarios from two different sectors were presented to the subjects with varying incoming traffic sequences. This is to avoid scenario recognition during the course of the experiment.

Analysis with regards to subject's behavior and workload rating were initially conducted to observe whether both sectors represent two sectors of different complexity, which would enable cross-sector transferability investigation on sector complexity measures. It is gathered that both sector indeed represent different levels of complexity, based on significant differences gathered from both subject's behavior and workload rating. It is also gathered that the number of aircraft present in a sector does not need to constitute the main factor that determines controller workload. Other sector complexity influencing variables, such as sector volume, route design and also geographical location of intercept points also contribute to the effect on how much effort was needed to control the sector. This is consistent with the concept of having different maximum number of aircraft *per* sector basis.

Initial correlation analysis were conducted to compare the SSD metric and unweighted NASA DD metric towards workload rating. The analysis is aimed at having a neutral comparison between both unweighted NASA DD and the SSD metric without the influence of any post-processing procedures. It is observed that based on initial correlation analysis, SSD is shown to have a higher level of correlation than unweighted NASA DD metric and number of aircraft. This is found in analysis based on both different sector and group of controllers.

Weighted NASA DD metrics from a collection of significant DV coupled together with weighting coefficient were gathered through regression analysis. Different sets of DV used to construct NASA DD metric for different sector and group controllers, were an indication of differences in controller's strategy in handling traffic within a sector. Thus, controller's individual differences would highly influence the construction of the DD metric. An improved correlation between weighted NASA DD and workload rating were gathered compared to unweighted NASA DD. However, when compared to SSD metric, only some weighted NASA DD metric showed a better correlation than SSD metric with workload rating.

It has been observed that when transferring a certain NASA DD model to a different sector or group of controllers, it has resulted in the metric not delivering the same level of correlation as previously found. The cross-sector and cross-group analysis also reveal that both NASA DD metric is sensitive towards different sector and group of controllers.

The original NASA DD metric was constructed based on a 3D airspace model with traffic samples from 36 high and low sectors, respectively. Due to the extent of data used in producing the metric, it is assumed that the NASA DD metric should be robust enough to be used on other traffic samples. However, the fact that the linear regression analysis to produce the weighted NASA DD metric in this experiment was gathered based on 2D airspace model using limited number of participants over a large number of variables, there could always be a possibility of the model being overfitted and in the end produce a poorer predictive performance. Exaggeration of minor fluctuations in the data could have deteriorated the method's performance. Nevertheless, the NASA DD metric should not be too sensitive to a specific sample size and should perform well on any sector design or group of controllers.

6.5 Conclusion

This chapter presents the result of the investigation of whether the SSD indeed presents a more reliable and objective sector complexity measure as it managed to show the same level of correlation under various sector designs and group of controllers settings. Comparisons between proposed SSD metrics and other known sector complexity measures, namely the number of aircraft and DD were conducted. From the experiment, it is concluded that the proposed method indeed represents a reliable and objective sector complexity measure, which could function better than number of aircraft, unweighted NASA DD metric and in certain conditions, than the weighted NASA DD metric. The SSD metric, which can be use in real-time situation without any post-processing procedures also appeared to be less sensitive than the NASA DD metric, towards controller differences as to sector design.

7 Conclusions and Recommendations

7.1 Retrospective

This thesis proposed the use of a constraint-based approach, namely through the Solution Space Diagram (SSD), to obtain an objective measure of controller taskload that would be insensitive to inter-controller variability and would also be transferable across sectors. In its most succinct form, the SSD method aims to relate the area of available solution spaces (to resolve potential conflict situations) with Air Traffic Controller (ATCO) taskload, whereby it was hypothesized that a lack of solution options would lead to a higher taskload.

Numerous off-line simulations and real-time human-in-the-loop experiments have been conducted during the course of this study to validate this hypothesis. The SSD metric was initially assessed by its potential to capture sector complexity under varying conditions, such as the number of aircraft, the route structure, traffic crossing angles, sector layout, etc., and its correlation with subjective workload ratings in more dynamic scenarios. To investigate the insensitivity to inter-controller variability and the cross-sector transferability properties of the SSD metric, a comparison study was done with other commonly known complexity metrics, such as traffic density and the Dynamic Density (DD) metric developed by NASA.

The next sections give an overview of the results and present the future recommendations and conclusions.

SSD in static scenarios

To investigate the Solution Space Diagram (SSD) capability to capture 'sector complexity', various complexity measures with various settings were studied in Chapter 3 using a static, 2-Dimensional (2D) traffic environment. Specifically, the aircraft crossing angles, aircraft speeds, horizontal proximities, number of aircraft, and bunching streams were varied and analyzed by the observed changes in the SSD area percentages of the no-go areas.

The static simulation results clearly demonstrated that different airspace designs and traffic conditions affect the availability of possible control options. Aircraft horizontal proximities and crossing angles are two sector complexity variables that demonstrate notable effects on the SSD. Smaller proximity results in more area covered on the SSD. Whereas when observing an incoming aircraft, larger intercept angles result in a lesser area covered on the SSD. Although the SSD metric showed to have potential in capturing sector complexity, to be able to correlate it to controller workload human-in-the-loop experiments were needed.

SSD in merging scenarios

The behavior of the SSD metrics in a dynamic environment was first explored in Chapter 4 through a human-in-the-loop experiment featuring a merging task on two different groups of subjects, namely a student and an expert group. A total of six different scenarios of 20-minutes were constructed to investigate the dynamic properties of the SSD metric and its correlation to subjective workload ratings. Four sector design variables were looked into, namely: incoming aircraft proximity, the number of traffic streams, intercept angle, and the traffic mix.

It was gathered from the experiment that the hypotheses regarding aircraft proximities, number of incoming aircraft streams, traffic mixes and intercept angle cannot be confirmed. Comparisons cannot be made between the workload and the SSD area properties, as in each sector complexity factor investigation, significant results was gathered in either workload rating or SSD area properties, but never in both. For example, for different intercept angle scenarios, only SSD area properties showed significant results. Whereas for different number of streams scenarios, only the workload ratings showed significant results.

An insufficient sample size or artifacts in the scenario design may have represented an insufficient difference between two different levels of sector complexity and thus might have hampered the possibility of investigating the effect of different sector complexity design toward controller's workload rating and SSD area properties.

Despite the fact that both groups performed differently and had different control strategies, the workload ratings for both groups were found to have a higher correlation with the SSD area properties than with the number of aircraft. This demonstrates that even when the SSD metric was unable to measure single sector complexity factor with the current sample size, the SSD metric is, overall, the better workload predictor.

SSD in conflict detection scenarios

An Air Traffic Controller is not only responsible for the supervision of an efficient and orderly flow of air traffic, but also for the safety of all traffic in his assigned sector. In Chapter 5 it was investigated how the SSD metric would relate to a controller's conflict detection performance. The experiment featured two levels of traffic densities and a variety of intercept angles within one sector geometry. The goal of the experiment was to explore whether the sector complexity can be measured using the SSD and also whether there would exist a common SSD area 'pattern' where controllers would start to detect a conflict pair. Short scenarios of two minutes were constructed that enabled multiple repetitions of each experimental condition and that minimized the effects of controller-induced complexity.

Based on the experiment, it was found that a higher traffic density indeed, as expected, results in higher workload and a corresponding increase in the SSD area properties. However, there is no common relation that can directly associate workload rating and SSD area properties in varying intercept angles situations. It is concluded, based on the findings in this experiment, that intercept angle is an intricate matter to be investigated as a single sector complexity construct, in a situation where the difficulty of identifying conflicting aircraft pairs is not only influenced by controller behavior but also by the neighboring traffic within the sector.

The experiment, however, did not show a clear threshold on SSD area percentage where a controller would start to detect a conflict pair. Thus, it is concluded that the SSD does not represent a trigger for conflict detection.

Sector complexity measure: A comparison

Although more sector complexity constructs can be investigated to verify that indeed the SSD could represent an objective measure of sector complexity, to conclude the research, a comparison study was done in Chapter 6 between the SSD metric, the NASA Dynamic Density (DD) metric, and the number of aircraft. In particular, these metrics were compared in terms of their sensitivity and transferability properties between two groups of participants and across two different sectors. The experiment featured two different sector geometries with a set of traffic scenarios that required controller interactions.

Based on the correlation results, the raw SSD metric showed to have the highest correlation with the subjective workload for both the two sectors and the two different groups of controllers, compared to the raw (unweighted) DD and the number of aircraft. Further investigations showed that the correlation of the weighted NASA DD metric with workload improved significantly after linear regression analysis. However, this can only be achieved through post-processing, which makes it vulnerable to reflect either a specific sector and/or a specific group of controllers. After linear regression analysis was conducted, the weighted NASA DD metric showed to became highly sensitive towards differences in sectors and groups of controllers. Based on these results, the SSD metric was found to be better able to correlate with workload and also to better withstand (i.e., be more robust to) sector and controller changes than the DD metric and the number of aircraft.

However, the fact that the linear regression analysis to produce the weighted NASA DD metric in this experiment was gathered based on 2D airspace model using limited number of participants over a large number of variables, there could always be a possibility of the model being overfitted and in the end produce a poorer predictive performance. Exaggeration of minor fluctuations in the data could have deteriorated the method's performance. Nevertheless, the NASA DD metric should not be too sensitive to a specific sample size and should perform well on any sector design or group of controllers.

7.2 Is the SSD a good sector complexity measure and workload estimator?

This research aimed at answering the question whether the SSD would represent an objective measure of sector complexity and a viable subjective workload predictor. By looking back at the results of the numerous off-line and real-time human-in-the-loop experiments, the short answer would be yes. However, this answer can only be 'yes' in the light of this research, where many assumptions, concessions, and simplifications needed to be made. These could have influenced the results and perhaps made the SSD metric appear to be overly promising. Before one could extrapolate the research findings into a real operational setting, several insights about some of the most important issues are discussed first.

First of all, the assumption that all en-route traffic were situated on one fixed flight level made the simulations in some aspects artificial and not very realistic. In practice this meant that controllers could only issue heading and/or speed clearances to direct aircraft and to keep them safely separated. In reality, controllers prefer to use flight level changes for en-route traffic. Speed changes for en-route traffic are quite rare given the narrow flight envelope at high altitudes. However, the controllers were able to control traffic in the simulations and the scenarios were designed in such a way that the omission of altitude control would not cause a significant negative effect on the measurements. However, there is another and perhaps more important consequence of considering 2D scenarios. That is, the NASA DD metric has been developed for three-dimensional (3D) traffic situations. As a result, the 2D nature of the simulations might have under-represented the value of the DD metric in terms of its correlation to workload. As such, it is unclear how the SSD metric would compare to the DD metric in 3D traffic environments. Before we can speculate about the possible results, it is first necessary to develop a 3D SSD metric (Zhou, 2011).

Second, and related to the previous assumption, the simplifications and concessions in terms of 'complexity' could have created artifacts in the simulations and scenarios. Airspace complexity is defined by both structural and flow characteristics of the airspace (Sridhar et al., 1998). It is an intricate subject and each complexity parameter is inter-related to one another, making it difficult to investigate the effect of a single parameter while not causing another parameter to change.

One negative effect of trying to isolate specific complexity parameters became clear from the conflict detection experiment described in Chapter 5. To mitigate the influence of aircraft speed variations on the perceived workload and safety performance, all aircraft in the scenarios had the same speed. In hindsight, this turned the conflict detection task into a relatively simple geometrical problem for the experienced controller. This way the scenarios might have become too easy and thus influenced the workload ratings. Thus, it is not unthinkable that a similar artifact emerged by fixating other complexity parameters in the other experiments. Another artifact occasionally has arisen in the dynamic interaction experiments. That is, controllers sometimes immediately interacted with aircraft as soon as they entered the sector. This changed the original designed conflict angles between aircraft and thus influenced the complexity. Thus, observing the effect of designed intercept angles (static feature), has become more difficult, as a result of human control behavior (dynamic feature).

Third, the sense of (a lack of) simulator realism might have affected the subjective workload ratings and strategies. Because the controllers participated in an experiment where the simulator only resembled a small portion of their work, the psychological relevance of the workload is questionable. Additionally, there were no real detrimental or punishable consequences for their actions, so controllers were bolder in trying out new strategies. It also affected the sense of danger and stress in controlling traffic. That is, even if controllers failed to maintain separation, it will only affect their performance during the experiment, but no lives were at stake. The experiments also assumed fast and identical responses to controller commands. This might have also changed the controller's usual behavior as it may have triggered an intentionally delayed command to resolve a traffic conflict. However, a sense of time pressure was imposed by running the traffic samples at four times speed.

Within the experiment assumptions, concessions, and simplifications, the 2D SSD metric has shown an encouraging potential as an objective measure of sector complexity and predictor of workload. The metric managed to perform better or at least at the same level with the number of aircraft and unweighted NASA DD metric. However, based on the results gathered so far, the metric still has its drawback when used as a measure of single sector complexity variable. This is mainly due to the difficulty in presenting a single sector complexity variable by means of human-in-the-loop dynamic experiment; and also the SSD might be better at depicting certain sector complexity variables (for example, proximity) than others (for example number of streams).

7.3 Recommendations for future research

Research and development in other areas, beyond the scope of this thesis, is needed to eventually mature the SSD as an objective measure of sector complexity and as a workload predictor in real operational environments. Apart from further exploring the broader abstraction of sector complexity constructs, the work in this thesis has led to the following recommendations.

Solution Space in the Third Dimension

The possibility of implementing the SSD in a 3-Dimensional (3D) problem is not far to reach. Initial concept studies have been conducted on an analytical 3D SSD (Zhou, 2011) and an interface-based 3D SSD (Lodder et al., 2011).

In the analytical solution, the 3D SSD area for the observed aircraft (A_{obs}) is comprised of two intersecting circles (both from the top and the bottom of the protected area) and the flight envelope of the controlled aircraft (A_{con}) comprising the rotation of the performance envelope around its vertical axis with 360°, resulting in a donut-shaped solution space. A simplified diagram of the solution space constructed by the protected area of the observed aircraft and the flight envelope of the controlled aircraft is illustrated in Figure 7.1. Further studies need to be conducted to verify

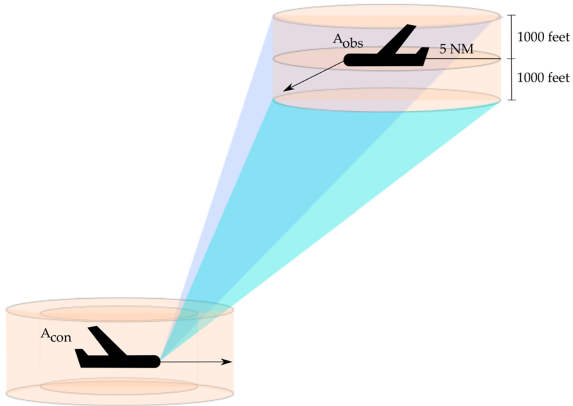


Figure 7.1: Two aircraft in 3D conditions.

the capability of the 3D SSD in efficiently measuring workload or sector complexity.

In a different study, the altitude dimension was integrated into a 2D-based SSD Air Traffic Controller (ATCO) display by Lodder et al. (2011). The altitude-extended SSD was calculated by filtering the intruder aircraft in accordance to their altitude relevance bands and cut off the SSD conflict zones by the slowest and fastest possible climb and descent profiles. In this way, the algorithm can discard conflict zones that can never lead to a conflict. Based on this algorithm, a display prototype has been developed that is able to show the effect of altitude changes to the controller. This display will be used in the future to perform human-in-the-loop experiments to assess the benefits of including altitude information in the 2D SSD ATCO displays.

As the current SSD also has the capability to present intent information, extension of the SSD metric to the third dimension will represent the metric as not only a 3D, but also a four-dimensional (4D) solution. Previous work by Hermes et al. (2009), d'Engelbronner et al. (2010) and Mercado Velasco et al. (2010) have introduced several methods to represent the SSD, while including the intent or trajectory information. Hermes et al. (2009) introduced the construction of intent-based SSD using combination of route segments to represent a known intent or specified routing. d'Engelbronner et al. (2010) on the other hand, propose a solution which approximates curved trajectories as a series of straight paths, that in the end resulting in rectilinear SSD no-go areas. Finally, Mercado Velasco (2009) has proposed a closed mathematical representation that represent the constraints imposed by proximate moving obstacles when the intended trajectories are know or can be estimated. The latter has the advantage of requiring low computational effort and is therefore suitable for real-time use.

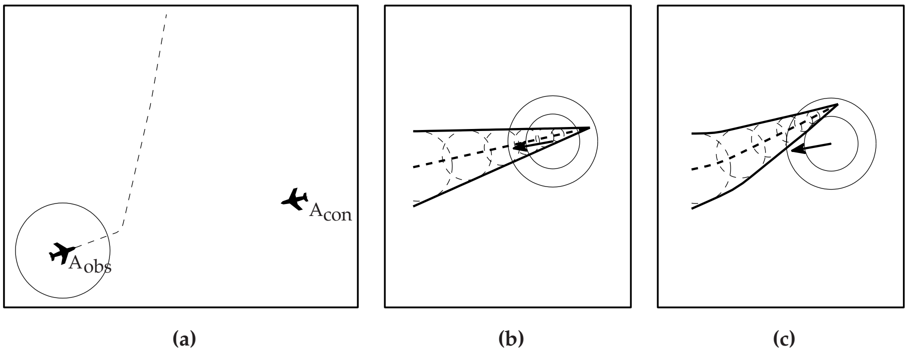


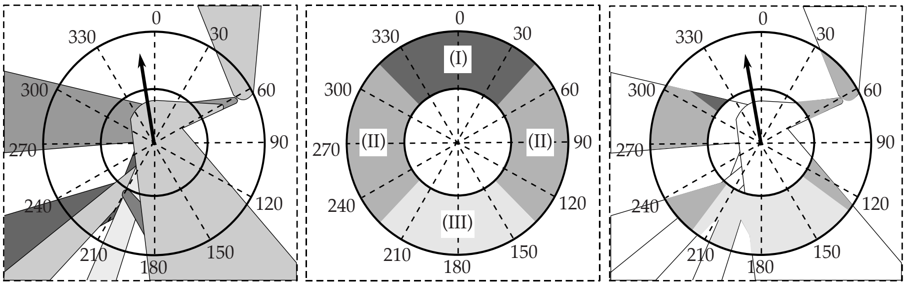
Figure 7.2: Effects of including intent information on the SSD. (a) A_{obs} sharing intent information with A_{con}. (b) SSD for A_{con} without the intent information of A_{obs}. (c) SSD for A_{con} with the intent information of A_{obs}.

Changes that intent information induces on the SSD are illustrated in Figure 7.2. Note that what would initially be regarded as a conflict under the SSD without intent information (Figure 7.2b), is no longer so when intent information is incorporated (Figure 7.2c). When this solution is coupled with the third dimension extension, the 4D solution will further meet future Air Traffic Management (ATM) constraints, which is likely to become a 4D trajectory-based system (Klomp et al., 2012).

Solution Space Observation Angle 'Weight'

The Solution Space presents 360° velocity vector options that leads to future separation violation for controlled aircraft in a sector. For pilots, the 180° in the current direction of the aircraft are considered to be the most important, as it gives all options that lead to the direction of destination. For controllers, 360° velocity vector options might still be applicable as it gives all possible options for controllers to change aircraft direction to maintain safe separation. However, taking a certain observation angle into account, or the weight of the SSD areas based on quadrants might improve the correlation with workload, as controllers similar to pilots are not likely to direct aircraft away from their original destination.

It is then assumed that angles in the direction of the current and future headings are more important than the opposite direction. This is also confirmed through findings in Chapter 5 where it was highlighted that certain observation angle limits were more significant for a certain task. To demonstrate this, a preliminary observation angle weighting technique was implemented where weighting factors were assigned to certain headings. An example situation, together with an application of an initial



(a) (b) (c)

Figure 7.3: Example of weight assignments to the SSD.

weighing factor is illustrated in Figure 7.3. In this example, the initial SSD has unsafe area (A_{whole}) of 46% (Figure 7.3a).

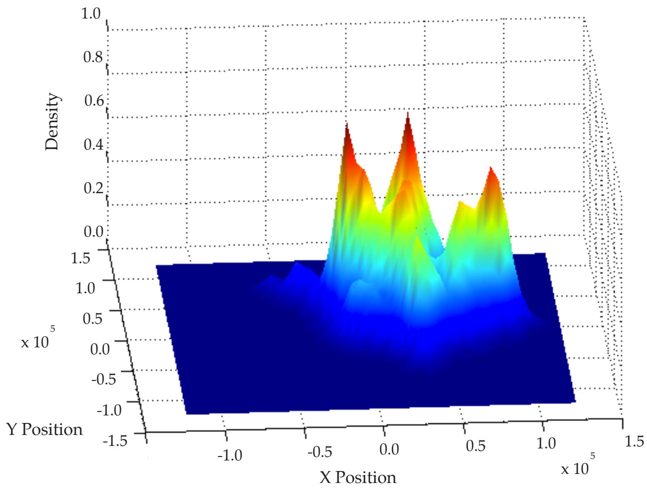
The observation angle is defined as the semi-sided angle relative to the velocity vector. For example, an observation angle of 90° means that half of the Solution Space has been taken into account. With that in mind, the SSD can be divided into 4 regions as seen in Figure 7.3b. The initial weighting value for each region is assigned as:

I) 100% of the SSD area value for area in current heading direction.
II) 75% of the SSD area value for area on both side of the velocity vector.
III) 50% of the SSD area value for area opposite the direction of the velocity vector.

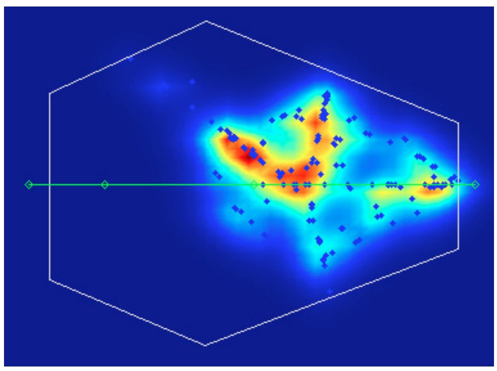
Based on this calculation method, an example of the same SSD properties as in Figure 7.3a is transformed into different heading factors as shown in Figure 7.3c. As a result, the mentioned aircraft will now have an A_{whole} of 29%. The idea here is that the SSD, which was covered by almost half of its possible safe area, might not contribute on the difficulty of a sector to that extent when almost all of its no-go areas are not in the aircraft current direction, and are therefore less relevant. It must be mentioned here, however, that the validity of this weighting method has not yet been tested. This will be a topic for future investigations.

Solution Space as Planning Tool

It is envisioned that the method can also be used as an offline tool in sector planning. This enables a more dynamic airspace sectorization or staff-planning than using the conventional maximum number-of-aircraft limit that is primarily driven by the air traffic controller's ability to monitor and provide separation, communication, and flow-control services to the aircraft in the sector. Other than using the SSD as a sector planning aid, the SSD may also be used as an operational tool. It is anticipated that by using the SSD as a display, controllers will have an additional visual assistance to navigate aircraft within the airspace. The SSD can serve as a collision avoidance tool (Mercado Velasco et al., 2010) or also a support tool for ATCOs, to indicate sector bottlenecks and hotspots. ATCOs.



(a) 3-Dimensional SSD map.



(b) 2-Dimensional SSD map.

Figure 7.4: Example of using the SSD covered area as a planning tool for highlighting hotspots.

Figure 7.4 illustrates the possibility of mapping hotspots using the SSD area properties. The mapping of the SSD is done based on the area properties known for each aircraft that fly through the sector. The SSD area properties are then matched to the locations of the aircraft and result in a 3D density plot as seen in Figure 7.4a. The resulting 3D plot is then transformed to a 2D map as seen in Figure 7.4b. The complexity map is constructed based on the notion that the blue regions represent regions with a low SSD area and the red regions represent regions with a high SSD area. It is anticipated that the mapping concept could help pinpoint the problematic or high SSD area percentage that later could be used as a planning help tool or real-time advisory tool for

Solution Space in Adaptive Automation

It is envisioned that in the future the Level of Automation (LOA) support for ATCOs is inferred not based on measured operator state, but through objective measures of sector complexity. An example of advanced automation that aims to balance operator workload demand between underload and overload is Adaptive Automation (AA). AA has been defined as the dynamic allocation of control over a system functions to a human operator and/or computer over time with the motivation of optimizing overall system performance (Rouse, 1977, Parasuraman, 1987, Scerbo, 1996, Kaber & Riley, 1999). Consequently, a more objective measure of sector complexity is needed to determine the level of task demand load imposed on the controller. This can be done using the SSD under varying traffic complexity.

Apart from using the SSD to better understand controller's workload, there are two other possible uses of the SSD metric within AA. Firstly, it can be directly used as a trigger switch for LOA support and secondly, it can be used in conjunction with physiological measurements to provide context information. In both solutions the SSD metric will provide either a measurement of sector complexity level, or function as a means for display to provide meaningful information to controllers.

7.4 Conclusions

This thesis aimed at investigating whether the SSD area of a 2D Air Traffic Control (ATC) separation problem can be used to assess the sector complexity and ATCO workload more accurately and objectively than other current metrics. An approach that is based on the behavior of 'zone of conflict' or 'no-go' velocity vector areas of neighboring aircraft has been put forward. The representations of no-go areas are essentially based on the constraints that limit the air traffic controller's decisions and actions within the aircraft performance limit. The more area covered on the solution space, that is, the fewer options the controller has to resolve conflicts, the higher will be the workload experienced by the controller.

Overall, the SSD has shown its capability in assessing the inherent difficulty of ATC situations. The correlation between the SSD metrics and workload ratings were found to be at least at the same level or better than the number of aircraft metric and the unweighted NASA DD metric. In some cases, the SSD even showed a higher correlation than the weighted NASA DD metric.

Initial quantitative analysis has shown that changes in the sector design variables such as traffic horizontal proximity, speed differences, intercept angle, traffic density, and traffic patterns can be illustrated through the changes in the SSD area properties. The thesis also addressed the effects of different sector complexity variables towards the SSD and workload ratings with human-in-the-loop experiment. Both measures were compared to see whether SSD metrics could function as a workload predictor. It was found that workload as a result of different sector complexity constructs can be depicted through the SSD, only in the case of different traffic density.

However, constructing different levels of complexity for example for different numbers of streams and intercept angles has been a challenging task. This is mainly due to how sector complexity variables are inter-related to each other. When changes were implemented to portray different intercept angles or number of streams, other sector complexity variables would also change. For example, increasing the intercept angle would mean larger distances between aircraft, when the initial Time To Conflict (TTC) is maintained, or a smaller TTC if the initial distance would be maintained. Also, changes in the number of streams within a sector may also contribute to a change in the aircraft horizontal proximities, when the initial airspace density would be maintained. Thus, the effect of investigating one sector complexity construct might be overshadowed by unintentional changes in other sector complexity constructs.

It is also found that there was no common SSD area percentage where controllers would start to detect conflict pairs. So, whereas the SSD performs well as an overall sector complexity measure, no particular SSD area patterns that trigger conflict identification, could be identified.

In spite of that, the SSD metric has shown to be a reliable metric which maintains its performance even when investigated using different groups of controllers with varying knowledge and experience on ATC. The SSD metric also has the capability to objectively measure sector complexity. The metric is found to be less sensitive to inter-controller variability and would also be better transferable across sectors than the weighted NASA DD metric.

It should be noted, however, that these results were gathered with regards to specific assumptions and experiment settings. To prove that the constraint-based method using the SSD metric is the most suited metric in measuring sector complexity construct in a real operational setting, a more extensive research regarding its performance and robustness should be done.

A

Obstacle Detection in Motion Planning

A.1 The History of Obstacle Representation

The initial introduction of an obstacle representation in maritime navigation using velocity vectors was identified as early as 1892 in a device called the Battenberg Course Indicator that was invented by Prince Louis of Battenberg. The Battenberg Course Indicator is a mechanical calculating device that is used in maritime navigation for determining the relative course and speed of other vessels compared to the user's own ship and was of use particularly when moving in large convoys (National Maritime Museum, London, n.d.). The same fundamental principles are still present in the Radar Navigation and Maneuvering Board Manual that was published in the *Radar Navigation and Maneuvering Board Manual, 7th Edition* (2001).

A.2 Development in Motion Planning

The underlying collision avoidance algorithm used in the Maneuvering Board since 1903 was currently also vastly used in areas of robotics, motion planning and aerospace. The method was used in a robotics and motion planning as a maneuvering-board approach in L. Tychonievich et al. (1989), velocity obstacle in Fiorini & Shiller (1993) and collision cone in Chakravarthy & Ghose (1998), to name a few.

The Maneuvering-board Approach

L. Tychonievich et al. (1989) first presented this approach to solve a 2-Dimension (2D) path planning problem involving moving obstacles. The maneuvering board mechanism, represents a simple and efficient approach to the problem of finding a collision-free path through a field of moving obstacles. By rejecting all potential velocity vectors for the vehicle which lie inside of an obstacle avoidance cone, a collision-free path is guaranteed for the controlled vehicle. The basic maneuvering board mechanism is illustrated in Figure A.1 for both static and moving obstacle.

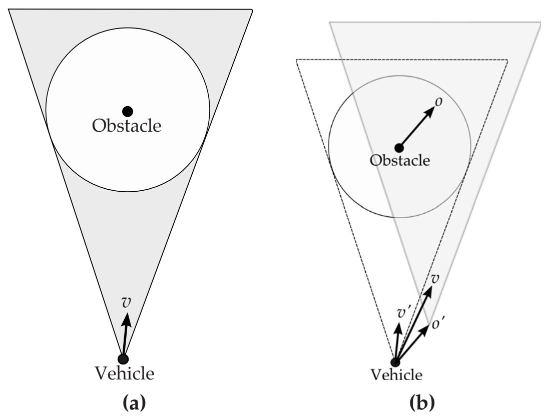


Figure A.1: Maneuvering Board method. (a) Observing stationary obstacle. (b) Observing moving obstacle.

In order to illustrate the avoidance cone diagram, it is assumed that the vehicle under control has perfect knowledge of the position and velocity of all objects within its vicinity. For example, Figure A.1 illustrates the avoidance cone diagram for two vehicles where the obstacle was represented as a circle that might represent the counter detection zone of a vessel located at the center and the controlled vehicle was represented as a point.

Figure A.1a shows the avoidance cone for a stationary obstacle. The cone is defined by the two tangent lines from the vehicle to the circle representing the obstacle. Any velocity vector, v which falls within this cone will put the vehicle on a collision course with the obstacle. The avoidance cone for a moving obstacle is constructed based on the same fundamental principle of using tangent lines from the vehicle to either side of the circle. However, the cone is now translated to the tip of the translated velocity vector of the obstacle (o'). Again, any velocity vector v for the vehicle that lies inside this cone represents a collision course with the moving obstacle.

In a later research by L. A. Tychonievich (2008), the set of velocities permitted by the maneuvering board algorithm for several aircraft condition can be graphically shown in Figure A.2. Using the plot, the velocities that would lead to a collision would be visible. This is the core of the maneuvering board algorithm.

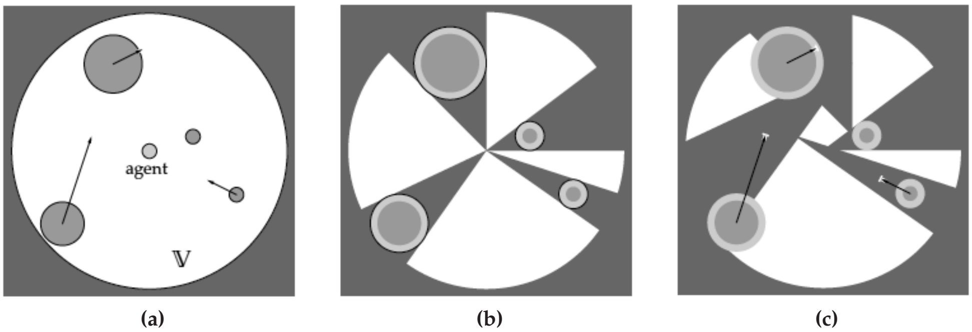


Figure A.2: Maneuvering Board graphical representation of four intruders situation (L. A. Tychonievich, 2008).

Figure A.2a illustrates the controlled agent and four obstacles together with a set of attainable velocities, v. The obstacle is then padded with the radius of the controlled agent. The avoidance cone of each motionless obstacle are illustrated and can be seen in Figure A.2b. By displacing the obstacle velocities, the final diagram can be gathered as seen in Figure A.2c. The white region within the diagram represents the safe areas for the agent to move, in order to avoid the obstacles in the area.

The Velocity Obstacle Approach

Fiorini & Shiller (1993) in their research, has developed the Velocity Obstacle (VO) method, which have the same basic principle as used in the maneuvering board approach. This method also illustrates each obstacle as a cone shaped forbidden velocity which represent the set of all velocities that will result in a collision at some moment in time, assuming that the observed object maintains its current velocity. If the object's velocity enters such cone, a collision would occur in a latter time. An example of a VO diagram between two objects, A and B is illustrated in Figure A.3.

Fiorini & Shiller (1998) also introduced Reachable Avoidance Velocities (RAV) which indicates the set of velocities that can be assigned to the controlled robot. For example, in the case as shown in Figure A.3 the only velocity vectors that is not suitable for the robot A to maneuver, based on the RAV are those within the area in

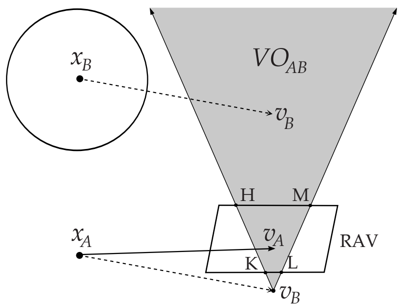


Figure A.3: Velocity obstacle method.

the segment H-M-L-K.

Latter studies on motion planning using the VO concept have showed examples of the VO usage in a 3-Dimensional (3D) environment (Fiorini, 1995) and also in dynamic environments (Fiorini & Shiller, 1998). In the research by Shiller et al. (2001), a dynamic representation of VO using the Non-Linear Velocity Obstacle (NLVO) was introduced. The NLVO were implemented in an interactive real-time simulation of a robot avoiding moving circular obstacles. Figure A.4 shows three obstacles and their NLVO, where two obstacles are seen to perform circular motion while another performs a straight constants motion as seen by the observer.

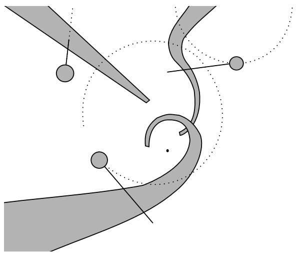


Figure A.4: Velocity obstacle graphical representation of dynamic motions (Shiller et al., 2001).

Advancement to the method based on velocity vectors is continuously researched in order to achieve a better collision detection and avoidance technique. Examples of researches in the area of obstacle detection in motion planning are summarized in Table A.1.

Table A.1: Methods using Velocity Vector approaches.

Method	Note
Collision Cone (CC) (Chakravarthy & Ghose, 1998)	The CC approach is proposed as an aid to collision detection and avoidance between irregularly shaped moving objects with unknown trajectories. Using this method, collision between a robot and an obstacle, where they both can be of any arbitrary shape and is moving in a dynamic environment can be detected.
Common Velocity Obstacle (CVO) (Yasuaki & Yoshiki, 2001)	The method expresses the VO using a CVO map in a Cooperative Collision Avoidance (CCA) task. Vehicles with future collision possibility will solve the problem using the CVO map, thus avoiding cancellation of actions between vehicles during collision resolution.
Non-Linear Velocity Obstacle (NLVO) (Shiller et al., 2001)	Introduced the non-linear velocity obstacle, which takes into account the shape, velocity and path curvature of the moving obstacle. Large et al. (2002) continued the work in building a complete autonomous navigation module and discuss the implementation in real-time situation.
Probabilistic Velocity Obstacle (PVO) (Kluge & Prassler, 2004)	A probabilistic extension to the velocity obstacle approach is used as a means for navigation and modeling uncertainty about the moving obstacle's decisions.
Velocity Space (Owen & Montano, 2005)	The method used the concept of estimated arriving time to compute the times to potential collision and potential escape. The dynamic environment is then mapped into the velocity space. The best motion command which satisfies an optimization criterion (typically the minimum time or the shortest path) is directly treated in the velocity space.
Reciprocal Velocity Obstacles (RVO) (van den Berg et al., 2008)	The approach takes into account the reactive behavior of the other agents by implicitly assuming that the other agents make a similar collision avoidance reasoning. Instead of choosing a new velocity for each agent that is outside the other agents velocity obstacle, the method chooses a new velocity that is the average of its current velocity and a velocity that lies outside the other agents velocity obstacle.
Forbidden Velocity Map (Damas & Santos-Victor, 2009)	Forbidden velocity region is mapped based on the possible collision region if the robot maintain the same direction of motion.
Generalized Velocity Obstacles (GVO) (Wilkie et al., 2009)	The approach generalizes the concept of velocity obstacles, which have been used for navigation among dynamic obstacles, and takes into account the constraints of a car-like robot.

B

Solution Space Diagram Plotter

To understand more on the behavior of Solution Space Diagram (SSD) regarding the aircraft position, speed and heading, a SSD plotter was developed. The plotter enables the user to study the effect of traffic situation on the SSD. The plotter consists of the Plan View Display (PVD) on the left side of the screen, three SSDs of different aircraft on the right side of the screen and also a number of buttons to navigate the controlled aircraft in the middle of the screen. The initial window setting with two aircraft present can be seen in Figure B.1. Any speed and heading control can be done through the top SSD by clicking the desired heading or speed for the controlled aircraft.

To help in visualizing the minimum separation of 5 NM radius between aircraft, a 'Guide' button is available that will show the minimum separation circle for every aircraft, the separation value between the aircraft (in NM), the separation between both aircraft and the merge point which is labeled with diamond shape mark (in NM), the percentage area covered on the main SSD, the width of the SSD at the same distance from the tip of the SSD and the progression of time to conflict of the Forbidden Beam Zone (FBZ). Figure B.2 illustrates the 'Guide' button execution where all the above information was presented to the user. The guide button will help the user to understand more on the relationship between the position, heading and speed towards the SSD properties such as the FBZ transformation on the SSD, also the width of the beam and the percentage of area covered. The FBZ is portrayed on the PVD has the same characteristic as appeared on the SSD.

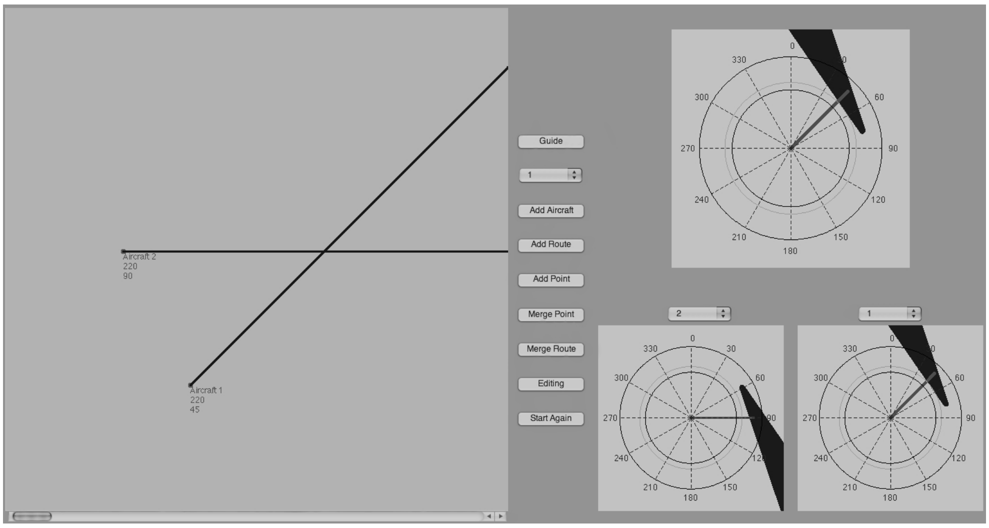


Figure B.1: Initial SSD plotter.

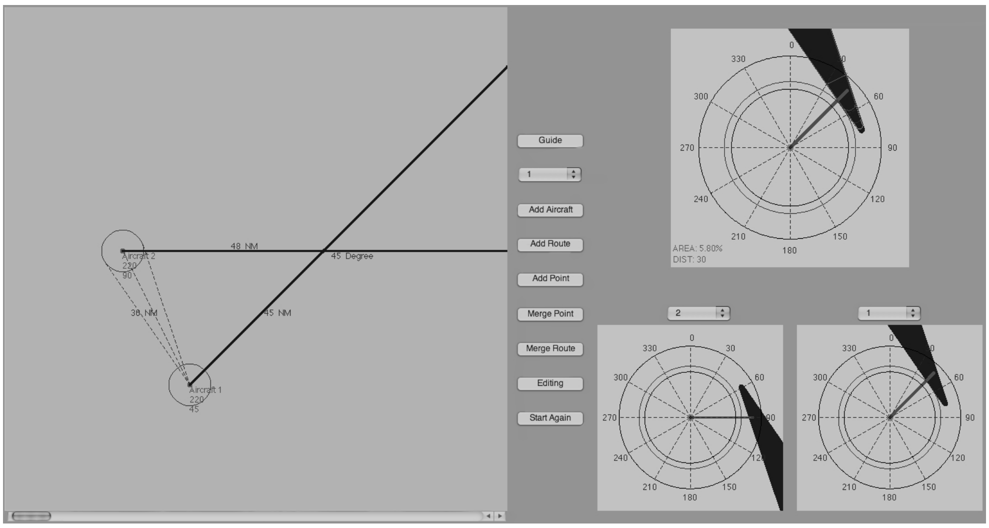


Figure B.2: SSD plotter with 'Guide' on.

The respective aircraft within the sector can be assigned to either a route or an overflight point. Both route and overflight point should be assigned beforehand. For a route, an unlimited number of route points can be assigned with a certain route speed. However for the overflight point, only one overflight point can be assigned with a certain speed and heading after reaching the point. The merging process can be monitored in Figures B.4a and B.4b for overflight point merge, and Figures B.5a and B.5b for route point merge. The initial three aircraft condition can be seen in Figure B.3.

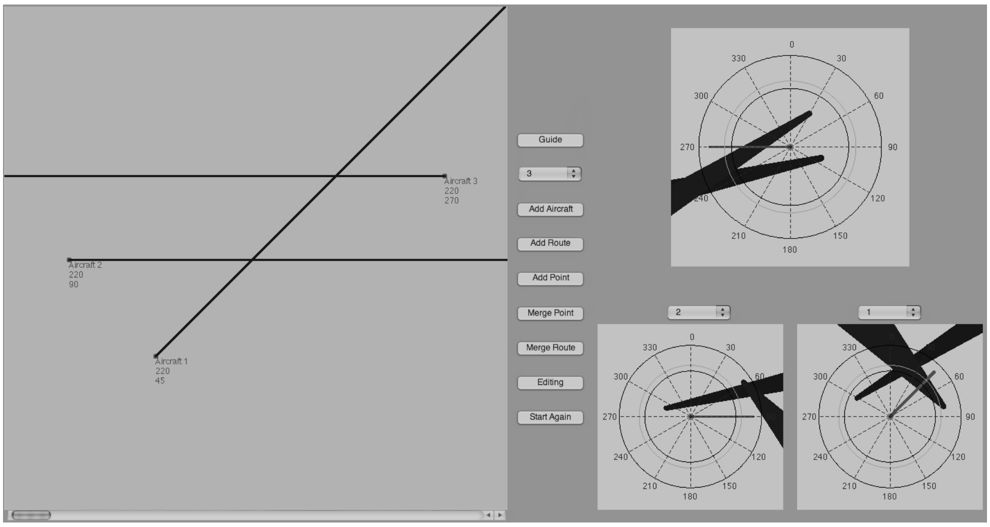
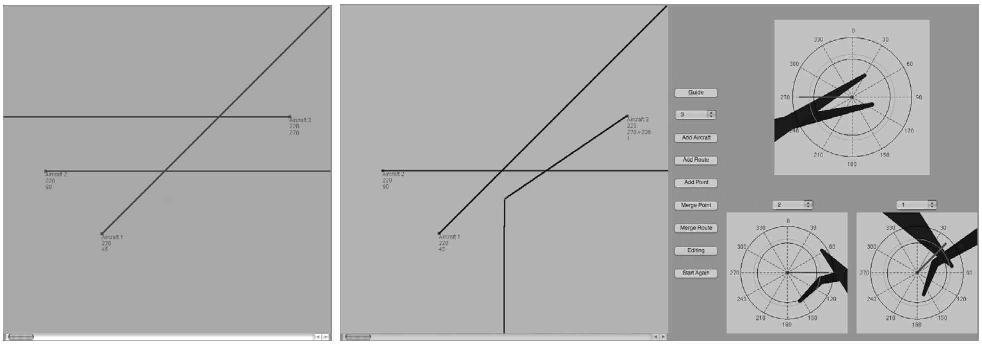


Figure B.3: Initial three aircraft condition.

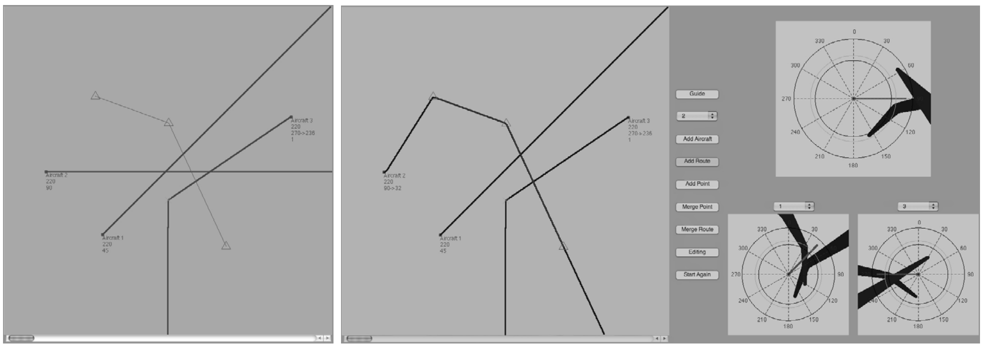
To initiate an overflight point merge, the point has to be assigned in the first place. Figure B.4a showed a point assigned on the sector. Aircraft 3 has been chosen to merge with the overflight point with speed of 210 knots and heading of 150° assigned after the aircraft has merged with the point. The SSD behavior towards Aircraft 1 and Aircraft 2 can then be observed from the respective SSD.

The same steps can also be done for the route assigning and route merging. Figure B.5a showed one example of route assignment where a route with three route points was added to the sector. Aircraft 2 is then merged to the route and the behavior of Aircraft 1 and Aircraft 3 can be observed from Figure B.5b in their respective SSD.



(a) **(b)**

Figure B.4: Example of situation with an overflight point. (a) Assigning an overflight point. (b) Merging with the overflight point.



(a) **(b)**

Figure B.5: Example of situation with route points. (a) Assigning a series of route points. (b) Merging with the assigned route.

The characteristic of the SSD of all the aircraft can be observed from the SSD window. If the timeline were moved, the progress of the FBZ on the SSD can be observed. Figure B.6 shows the progress of all aircraft after a certain moment in time in terms of position and also the SSD progress over time. As time progresses the positions of the aircraft get closer together and the increase in width of the SSD can be observed.

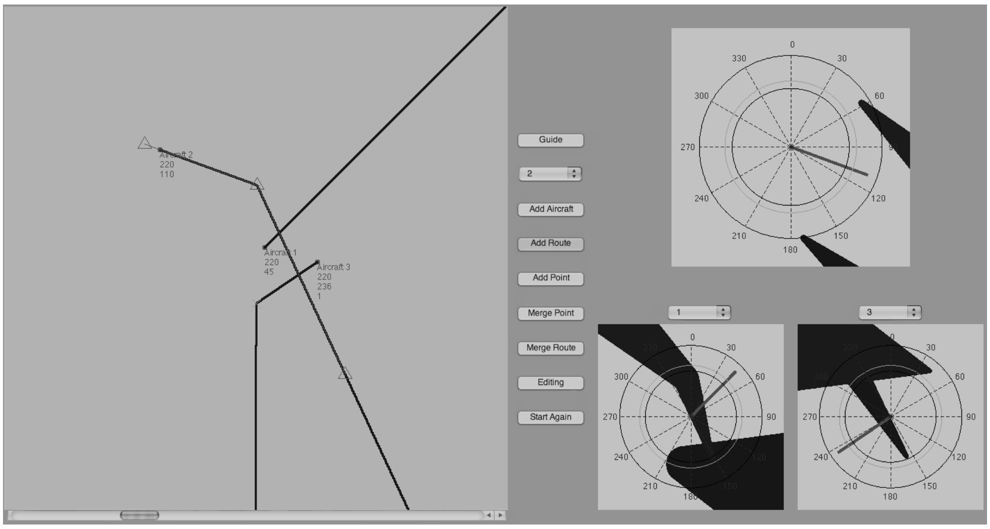


Figure B.6: Aircraft position and SSD progress over time.

Overall the reason behind developing the SSD plotter is to investigate the behavior of SSD in the act of merging a point or route and also the progression of the SSD in time. The behavior of the SSD in relation to position, speed and heading can also be investigated regarding the possibility of emerging patterns.

References

Airbus, S. A. S. (2012, December). *Global Market Forecast 2012-2031* (Tech. Rep.). Airbus.

Bilimoria, K. D. (2000, Aug). A Geometric Optimization Approach to Aircraft Conflict Resolution. In *AIAA Guidance, Navigation, and Control Conference.* Colorado, US.

Blom, H. A. P., & Klompstra, M. B. (2011). *iFly Publishable Final Activity Report* (Tech. Rep.). National Aerospace Laboratory.

Carbone, C., Ciniglio, U., Corraro, F., & Luongo, S. (2006, Dec). A Novel 3D Geometric Algorithm for Aircraft Autonomous Collision Avoidance. In *IEEE Conference on Decision and Control* (p. 1580-1585). CA, USA.

Chakravarthy, A., & Ghose, D. (1998, Sep). Obstacle Avoidance in a Dynamic Environment: A Collision Cone Approach. *IEEE, System, Man and Cybernetics - Part A: Systems and Humans*, *28*(5).

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

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

Damas, B., & Santos-Victor, J. (2009, Oct). Avoiding Moving Obstacles: The Forbidden Velocity Map. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 4393-4398). St. Louis, USA.

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

EASA. (2012). *Annual Safety Review 2012* (Tech. Rep.). Cologne, Germany: European Aviation Safety Agency.

Endsley, M. R., & Rodgers, M. D. (1998). Distribution of Attention, Situation Awareness, and Workload in a Passive Air Traffic Control Task: Implications of Operational Errors and Automation. *Air Traffic Control*, *6*(1), 21-44.

Eurocontrol. (2010, Aug). *Single European Sky ATM Research (SESAR) Joint Undertaking.* Available from `http://www.sesarju.eu/`

FAA. (2011, Mar). *NextGen Implementation Plan.* Available from `http://www.faa.gov/nextgen/media/ng2011_implementation_plan.pdf`

Farmer, E., Brownson, A., & QinetiQ. (2003, May). *Review of Workload Measurement, Analysis and Interpretation Methods* (Tech. Rep.). EUROCONTROL.

Fiorini, P. (1995). *Robot Motion Planning Among Moving Obstacles*. Unpublished doctoral dissertation, University of California.

Fiorini, P., & Shiller, Z. (1993). Motion Planning in Dynamic Environments Using the Relative Velocity Paradigm. In *IEEE International Conference of Automation and Robotics* (p. 550-566). Los Alamitos, CA.

Fiorini, P., & Shiller, Z. (1998, Jul). Motion Planning in Dynamic Environments Using Velocity Obstacles. *The International Journal of Robotics Research*, *17*, 760-772.

Goss, J., Rajvanshi, R., & Subbarao, K. (2004, Aug). Aircraft Conflict Detection and Resolution using Mixed Geometric and Collision Cone Approaches. In *AIAA Guidance, Navigation, and Control Conference and Exhibition.* Rhode Island, USA.

Hart, S. G., & Saveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Ed.), *Human Mental Workload* (p. 139-183). Amsterdam, North Holland.

Hermes, P., Mulder, M., van Paassen, M. M., Beoring, J. H. L., & Huisman, H. (2009, Nov). Solution Space-Based Analysis of Dificulty of Aircraft Merging Tasks. *Journal of Aircraft*, *46*(6), 1-21.

Hilburn, B. G. (2004). *Cognitive Complexity in Air Traffic Control - A Literature Review* (Tech. Rep. No. EEC Note 04/04). Bretigny-sur-Orge, France: EUROCONTROL.

Hilburn, B. G., & Jorna, P. G. A. M. (2001). Workload and Air Traffic Control. In P. A. Hancock and P. A. Desmond (Ed.), *Stress, Workload and Fatigue: Theory, Research and Practice* (p. 384). Hillsdale, New Jersey, USA: Erlbaum.

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

Kaber, D. B., & Riley, J. (1999). Adaptive Automation of a Dynamic Control Task Based on Secondary Task Workload Measurement. *International Journal of Cognitive Ergonomics*, *3*, 169-187.

Kimball, K. A., Hoffman, M. A., & Nossaman, R. O. (1973). Differential Velocity and Time Prediction of Motion. In *Perceptual & Motor Skills* (Vol. 36, p. 935-945).

Kinney, G. C., Spahn, J., & Amato, R. A. (1977). *The Human Element in Air Traffic Control: Observations and Analyses of the Performance of Controllers and Supervisors in Providing ATC Separation Services* (Tech. Rep. No. Report No. MTR-7655). McLean, VA: METREK Division of the MITRE Corporation.

Kirwan, B., Scaife, R., & Kennedy, R. (2001, Jun). Investigating Complexity Factors in UK Air Traffic Management. *Human Factors and Aerospace Safety*, *1*(2).

Klomp, R., van Paassen, M. M., Borst, C., & Mulder, M. (2012). Joint Human-Automation Cognition through a Shared Representation of 4D Trajectory Management. In *Proceedings of the SESAR Innovation Days* (p. 1-7).

Kluge, B., & Prassler, E. (2004, Apr). Reflective Navigation: Individual Behaviors and Group Behaviors. In *IEEE International Conference on Robotics and Automation* (p. 4172-4177). LA, USA.

Kopardekar, P., & Magyarits, S. (2002, Oct). Dynamic Density: Measuring and Predicting Sector Complexity. In *21st Digital Avionics System Conference* (p. 1-9). Pascataway, NJ.

Kopardekar, P., & Magyarits, S. (2003, Jun). Measurement and Prediction of Dynamic Density. In *5th ATM Seminar* (p. 1-10). Budapest, Hungary.

Koros, A., Rocco, P. D., Panjwani, G., Ingurgio, V., & D'Arcy, J.-F. (2004). *Complexity in Air Traffic Control Towers: A Field Study* (Technical Note DOT/FAA/CT-TN03/14). Springfield, Virginia: NTIS.

Large, F., Sekhavat, S., Shiller, Z., & Laugier, C. (2002, Dec). Using Non-Linear Velocity Obstacles to Plan Motions in a Dynamic Environment. In *7th International Conference on Control, Automation, Robotics and Vision* (p. 734-739). Singapore.

Laudeman, I. V., Shelden, S., Branstron, R., & Brasil, C. (1998, Apr). *Dynamic Density: An Air Traffic Management Metric* (Tech. Rep. No. NASA-TM-1998-112226). NASA Center for AeroSpace Information.

Lee, K., Feron, E., & Pritchett, A. R. (2009). Describing Airspace Complexity: Airspace Response to Disturbances. *Journal of Guidance, Control and Dynamics*, *32*(1), 210-222.

Leplat, J., & Bisseret, A. (1966). Analysis of the Processes Involved in the Treatment of Information by the Air Traffic Controller. In *Controller* (Vol. 1, p. 13-22).

Lillo, F., Pozzi, S., Ferrara, G., & Lieutaud, F. (2009, Dec). Coupling and Complexity of Interaction of STCA Networks. In *8th Innovative Research Workload Exhibition.* Bretigny-sur-Orge, France.

Lodder, J., Comans, J., van Paassen, M. M., & Mulder, M. (2011). Altitude-extended Solution Space Diagram for Air Traffic Controllers. In *International Symposium on Aviation Psychology.* Dayton, USA.

Majumdar, A., & Ochieng, W. (2007). Air traffic Control Complexity and Safety: A Framework for Sector Design Based Upon Controller Interviews of Complexity Factors. *Journal of the Transportation Research Board*, 70-80.

Majumdar, A., Ochieng, W., McAuley, G., Lenzi, J. M., & Lepadatu, C. (2004). The Factors Affecting Airspace Capacity in Europe: A Framework Methodology based on Cross Sectional Time-Series Analysis using Simulated Controller Workload Data. *Journal of Navigation*, *57*(3), 385-405.

Masalonis, A. J., Callaham, M. B., & Wanke, C. R. (2003, Jan). *Dynamic Density and Complexity Metrics for Realtime Traffic Flow Management* (Tech. Rep.). Center for Advanced Aviation System Development (CAASD), The MITRE Corporation.

Meckiff, C., Chone, R., & Nicolaon, J.-P. (1998, Dec). The Tactical Load Smoother for Multi-Sector Planning. In *2nd USA/Europe Air Traffic Management R&D Seminar.* Orlando, FL.

Mercado Velasco, G. A. (2009). *The Solution Space Diagram as Workload Alleviator in the ATC Task of Merging Aircraft*. Unpublished master's thesis, Faculty of Aerospace Engineering, Delft University of Technology.

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

Mogford, R. H., Guttman, J., Morrow, S. L., & Kopardekar, P. (1995, Jul). *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature* (Tech. Rep.). U.S Department of Transportation, Federal Aviation Administration.

Moray, N. (1977, Aug). Mental Workload - Its Theory and Measurement. In *Proceedings of the NATO Symposium on Theory and Measurement of Mental Workload.* New York: Plenum Press.

Moray, N. (1982). Subjective Mental Workload. *Human Factors*, *24*, 25-40.

National Maritime Museum, London. (n.d.). *Battenberg's Course Indicator Mark 5.* Retrieved 1 Oct 2013, from `http://collections.rmg.co.uk/collections/objects/201830.html`

Nunes, A., & Kirlik, A. (2005). An Empirical Study of Calibration in Air Traffic Control Expert Judgment. In *Human Factors and Ergonomics Society 49th Annual Meeting* (p. 422-426).

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload Assessment Methodology. In K. Boff & L. Kaufmann & J.Thomas (Ed.), *Handbook of perception and human performance* (Vol. 2, p. 42-29). New York: John Wiley.

Owen, E., & Montano, L. (2005, Aug). Motion Planning in Dynamic Environments Using Velocity Space. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 2833-2838). Alberta, Canada.

Parasuraman, R. (1987). Human-Computer Monitoring. *Human Factors*, *29*, 695-706.

Prandini, M., Piroddi, L., Puechmorel, S., & Brazdilova, S. L. (2011, Sep). Toward Air Traffic Complexity Assessment in New Generation Air Traffic Management Systems. *IEEE Transactions on Intelligent Transportation Systems Society*, *12*(3), 809-818.

Puechmorel, S., & Delahaye, D. (2009). New Trends in Air Traffic Complexity. In *Proceedings EIWAC.* Tokyo, Japan.

Radar Navigation and Maneuvering Board Manual, 7th Edition (Tech. Rep.). (2001). Maritime Safety Information Center, National Imagery and Mapping Agency.

Rantanen, E. M., & Nunes, A. (2005). Hierarchical Conflict Detection in Air Traffic Control. *The International Journal of Aviation Psychology*, *15*(4), 339-362.

Remington, R. W., Johnston, J. C., Ruthruff, E., Gold, M., & Romera, M. (2000). Visual Search in Complex Displays: Factors Affecting Conflict Detection by Air Traffic Controllers. *Human Factors*, *42*(3), 349-366.

Rodgers, M. D., Manning, C. A., & Kerr, C. S. (1994, Oct). Demonstration of POWER: Performance and Objective Workload Evaluation Research. In *Proceeding of the Human Factors and Ergonomics Society 38th Annual Meeting.* Nashville, TN.

Rouse, W. B. (1977). Human-Computer Interaction in Multi-Task Situations. *IEEE Transactions on Systems, Man and Cybernetics*, *7*, 384-392.

Scerbo, M. W. (1996). Theoretical Perspectives on Adaptive Automation. In (p. 37-64). Mahwah, NJ: Erlbaum.

Scerbo, M. W. (2007). Adaptive Automation. In R. Parasuraman & M. Rizzo (Ed.), *Neuroergonomics: The Brain at Work* (p. 239-252). Oxford University Press.

Schroeder, D. J. (1982). The Loss of Prescribed Separation Between Aircraft: How Does it Occur? In *Behavioral Objectives in Aviation Automated Systems Symposium* (p. 257-269). Washington, DC.

Shiller, Z., Large, F., Sekhavat, S., & Laugier, C. (2001, May). Motion Planning in Dynamic Environments: Obstacles Moving Along Arbitrary Trajectories. In *IEEE International Conference on Robotics and Automation* (p. 3716-3721). Seoul, Korea.

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

Tattersall, A., & Foord, P. (1996). An Experimental Evaluation of Instantaneous Self-Assessment as a Measure of Workload. *Ergonomics*, *39*(5), 740-748.

Tulga, M. K. (1978). *Dynamic Decision Making in Multitask Supervisory Control: Comparison of an Optimal Algorithm to Human Behaviour*. Unpublished doctoral dissertation, Cambridge, MIT Man Machine System Laboratory.

Tychonievich, L., Zaret, D., Mantegna, J., Evans, R., Muehle, E., & Martin, S. (1989, Aug). A Maneuvering-Board Approach to Path Planning with Moving Obstacles. In *International Joint Conference on Artificial Intelligence* (p. 1017-1021). Michigan, USA.

Tychonievich, L. A. (2008). *Simulation and Visualization of Environments with Multidimensional Time*. Unpublished master's thesis, Department of Computer Science, Brigham Young University.

Van Dam, S. B. J., Abeloos, A. L. M., Mulder, M., & van Paassen, M. M. (2004). Functional Presentation of Travel Opportunities in Flexible Use Airspace: an EID of an Airborne Conflict Support Tool. In *IEEE International Conference on Systems, Man and Cybernatics* (Vol. 1, p. 802-808).

van den Berg, J., Lin, M., & Manocha, D. (2008, May). Reciprocal Velocity Obstacles for Real-Time Multi-Agent Navigation. In *IEEE International Conference on Robotics and Automation* (p. 1928-1935). CA, USA.

Veltman, H. J. A., & Jansen, C. (2004). The Adaptive Operator. In D. A. Vincenzi & M. Mouloua & P. A. Hancock (Ed.), *Human Performance, Situation Awareness and Automation: Current Research and Trends* (Vol. II, p. 7-10). Mahwah, NJ: Erlbaum.

Whiteley, M. (1999, Nov). *PHARE Advanced Tools Tactical Load Smoother* (Tech. Rep.). European Organisation for the Safety of Air Navigation (EUROCONTROL).

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (Third ed.). Upper Saddle River, NJ: Prentice Hall.

Wilkie, D., Berg, J. van den, & Manocha, D. (2009, Oct). Generalized Velocity Obstacles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 5573-5578). St. Louis, USA.

Yasuaki, A., & Yoshiki, M. (2001, Nov). Collision Avoidance Method for Multiple Autonomous Mobile Agents by Implicit Cooperation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (p. 1207-1212). Hawaii, USA.

Zhou, W. (2011). *The 3D Solution Space: Metric to Assess Workload in Air Traffic Control*. Unpublished master's thesis, Department of Control and Simulation, Delft University of Technology.

Samenvatting

Methode voor de Bepaling van Sectorcomplexiteit door Middel van Evaluatie van de Beschikbare Oplossingsruimte

Siti Mariam binti Abdul Rahman

In het verleden zijn verschillende methoden geïntroduceerd met als doel het optimaliseren van het (luchtruim-)sectorontwerp en de toewijzing van luchtverkeersleiders aan sectoren. Deze optimalisatie wordt gedaan om groei te accommoderen, productiviteit te verhogen en met name om de veiligheid van het luchtverkeer te garanderen. Om dit te bereiken is een meer omvattend begrip van de menselijke werklast, in het bijzonder die van luchtverkeersleiders, vereist.

Binnen luchtverkeersleiding geldt een maximum aantal vliegtuigen per sector welke de luchtverkeersleider verondersteld wordt tegelijkertijd te kunnen aansturen. Dit maximum wordt bepaald op basis van experimenten en aan de hand van een schatting van subjectieve werklast, welke sector-eigen zijn. Deze grens dient niet overschreden te worden om een redelijke en aanhoudende mate van werklast te behouden. Echter, het bepalen van complexiteit op basis van het maximum aantal vliegtuigen houdt geen rekening met de dynamische aard van het luchtverkeer. Dit beperkt de mogelijkheid voor het accommoderen van groei. Om de strategische beslissingen, waarin de werkbelasting voor luchtverkeersleiding een rol speelt, beter te ondersteunen, is er een behoefte aan betere indicatoren dan uitsluitend het aantal vliegtuigen.

Grootheden, zoals bijvoorbeeld Dynamische Dichtheid (Eng: Dynamic Density, DD), zijn ontwikkeld en voorgesteld als maat voor sectorcomplexiteit. Deze maken gebruik van een gewogen combinatie van statische en dynamische eigenschappen van het luchtruim, zoals bijvoorbeeld het aantal vliegtuigen dat door een sector vliegt, de fractie van klimmende, kruisende en dalende vliegtuigen, en de horizontale afstand tussen vliegtuigen. De voorgestelde weegfactoren worden bepaald door regressieanalyses op basis van het oordeel van experts voor een bepaald sector ontwerp. Hierdoor zijn de grootheden sterk afhankelijk van zowel de sector als de individuele expert en daarom niet uniform toepasbaar voor een bredere groep gebruikers of sectorontwerpen. Een nauwkeurige ijking zou nodig zijn om zulke grootheden aan te passen naar elke individuele gebruiker en sector.

In een poging een meer objectieve indicator van sectorcomplexiteit en voorspeller van werklast te ontwikkelen, onderzoekt dit proefschrift een methode welke gebruik maakt van de beschikbare oplossingsruimte op basis van het oplossingsruimtediagram (Eng: Solution Space Diagram, SSD). In essentie is het SSD een weergave van de beperkingen en mogelijkheden tot manoeuvreren om luchtverkeersconflicten op te lossen door middel van zowel richting - als snelheidsinstructies. Het SSD kan beschreven worden als de beschikbare stuurruimte voor het vliegtuig, rekening houdend met andere waargenomen vliegtuigen in de nabijheid. De opbouw van het SSD is gebaseerd op de projectie van het veiligheidsgebied rondom het waargenomen vliegtuig, wat bepaald wordt door de minimale separatie van 5 zeemijl tussen vliegtuigen. Vanuit het perspectief van een bepaald vliegtuig introduceert elk ander toestel een no-go area - ook wel 'conflictgebied' genaamd - op het SSD. Het binnendringen van dit gebied wordt een conflict of separatieverlies genoemd.

Initïele studies naar gebruik van het SSD suggereerden dat deze methode inderdaad een mogelijke manier is om de dynamiek van de taaklast te bepalen, en in sommige gevallen zelfs in staat is om werklast te voorspellen. Het doel van dit proefschrift was om te onderzoeken of het gebruik van het op manoeuvreerbeperkingen gebaseerde SSD in staat is om de dynamiek van luchtverkeersleidingstaken op een objectieve en betrouwbare manier te bepalen. Dit zou het een nuttig middel maken voor de beoordeling van toekomstige luchtverkeersleidingsconcepten.

De belangrijkste hypothese van dit proefschrift stelt dat hoe meer oppervlakte beperkt is in de oplossingsruimte, met andere woorden hoe minder mogelijkheden de luchtverkeersleider heeft om conflicten op te lossen, hoe moeilijker de verkeersituatie is en hoe hoger daarom de, door de verkeersleider ervaren, werklast zal zijn.

In dit proefschrift worden hoofdzakelijk twee oppervlakteberekeningen gebruikt. Dit zijn de totale oppervlakte van het onveilig gebied (A_{whole}) voor een bepaald vliegtuig en het gemiddelde onveilige oppervlak (A_{mean}) - het gemiddelde van de oplossingsruimte van alle toestellen binnen de sector. Beide oppervlaktebepalingen worden gebruikt om het effect te begrijpen van verschillende metrieken van sector complexiteit op de beschikbare oplossingsruimte. De A_{whole} wordt berekend op basis van het totale bedekte oppervlak binnen de richtingsband tussen minimum - en maximumsnelheid - de prestatielimieten - van elk individueel toestel. De A_{mean} wordt vervolgens samengesteld uit de som van A_{whole} van alle vliegtuigen in de sector, gedeeld door het totaal aantal vliegtuigen in de sector. Waar A_{whole} de beperkingen op ieder individueel vliegtuig vertegenwoordigt, geeft A_{mean} een representatie voor de gehele sector.

Dit onderzoek is zo opgezet dat verschillende relevante verkeerssituaties of condities worden gegenereerd door middel van computersimulatie van variabele condities, of door evaluatie van menselijke prestaties en werklast van een luchtverkeersleidingstaak, door middel van experimenten met menselijke proefpersonen. Drie verschillende taken voor de verkeersleiding in de kruisvlucht worden onderzocht, te weten: Het samenvoegen van vliegtuigen op de route, het identificeren en oplossing van conflicten en het klaren van vliegtuigen naar routepunten. Doel hiervan is om de robuustheid en toepasbaarheid van de SSD metrieken te bepalen. In de experimenten met menselijke proefpersonen werd het SSD naderhand gebruikt om de complexiteit van de sector te berekenen. Vervolgens is de correlatie bepaald van de uitkomst van deze analyse met de experimenteel gemeten subjectieve werklastindicaties van de proefpersonen. Elk hoofdstuk in dit proefschrift beschrijft een onderzochte mogelijkheid om, door gebruik van de beschikbare oplossingsruimte, de complexiteit van een sector te bepalen en de werklast te voorspellen.

Het onderzoek begint met een analyse van de effecten van een aantal factoren in sectorcomplexiteit op de verschillende oppervlakte-eigenschappen van het SSD. Hoofdstuk 3 presenteert een analyse van een studie naar specifieke gevallen waarin twee vliegtuigen elkaar ontmoeten uit verschillende richtingen, met verschillende snelheden en met verschillende nog af te leggen lengtes van routes. De verschillen in deze parameters in sectorontwerp zijn vervolgens systematisch gerelateerd aan veranderingen in de eigenschappen van oppervlakken op het SSD. De horizontale nabijheid tussen vliegtuigen en de hoek waarmee vliegtuigen elkaar kruisen zijn twee voorbeelden van parameters die een waarneembaar effect hebben op het SSD. Vliegtuigen welke dichtbij zijn leiden tot een meer bedekt oppervlak op het SSD. Grotere kruisingshoeken leiden tot kleinere oppervlakken op het SSD.

In de studies met menselijke proefpersonen, welke zijn beschreven in Hoofdstukken 4 en 5, is aangetoond dat het SSD een beter - of tenminste gelijk - correlatieniveau heeft met subjectieve indicaties van werklast dan een metriek uitsluitend gebaseerd op het aantal vliegtuigen in de sector. In een poging om de mogelijkheid te onderzoeken om de werkbelasting van verschillende sectorcomplexiteit factoren te meten, zijn scenario's met varierende laterale nabijheid, kruisingshoeken, aantallen verkeersstromen en verkeersdichtheden gecreëerd en onderzocht in experimenten met menselijke proefpersonen. Wanneer de verkeersdichtheid varieerde, kon de verandering van werklast voorspeld worden door middel van grootheden gebaseerd op het SSD. Hogere verkeersdichtheid leidde tot een hogere werklast, wat ook gezien kon worden in de eigenschappen van de oppervlaktes binnen het SSD.

Echter, het ontwerpen van verschillende niveaus van complexiteit voor verschillende onderlinge afstanden, aantallen verkeersstromen en kruisingshoeken blijkt een uitdaging. Dit komt voornamelijk doordat het wijzigen van een bepaalde complexiteitsfactor leidt tot onbedoelde wijziging van andere factoren. Een verandering in het aantal verkeersstromen kan bijvoorbeeld tot een verandering in de onderlinge afstanden leiden indien de verkeersdichtheid wordt behouden. Bij scenario's met verschillende kruisingshoeken leidt een hogere kruisingshoek tot grotere onderlinge afstanden indien de tijd tot het conflict gelijk wordt gehouden. Indien de afstanden gelijk worden gehouden neemt de tijd tot het conflict af. Hierdoor is het mogelijk dat het onderzochte effect van een bepaalde factor van complexiteit overschaduwd wordt door onbedoelde veranderingen van andere factoren.

Ondanks deze beperking is aangetoond dat het SSD een betrouwbare indicator vormt, welke zelfs in staat is om haar prestaties te behouden bij verschillende groepen luchtverkeersleiders met aanzienlijke verschillen in kennis en ervaring. Om de toepasbaarheid en mogelijke voorbeelden van de SSD-grootheid beter in kaart te brengen wordt in Hoofdstuk 6 de indicator vergeleken met een breed geaccepteerde grootheid, de dynamische dichtheid DD. Uit het onderzoek blijkt dat het SSD in staat is de inherente moeilijkheidsgraad van een luchtverkeerssituatie in te schatten. Er is aangetoond dat de correlatie tussen de SSD metrieken en werklast gelijk of beter was dan uitsluitend het aantal vliegtuigen of de ongewogen DD grootheid. In een aantal gevallen was de correlatie zelf hoger dan die van de gewogen DD. De SSD grootheid heeft tevens de mogelijkheid om de sectorcomplexiteit objectief te bepalen waarbij is aangetoond dat deze dan minder gevoelig is voor verschillen tussen individuen, en tevens beter toepasbaar over meerdere sectoren, dan de gewogen DD indicator.

De resultaten van de verschillende computersimulaties en de experimenten met menselijke proefpersonen laten zien dat de voorgestelde SSD grootheid een veelbelovende mogelijkheid is voor het objectief bepalen van de sectorcomplexiteit en het voorspellen van werklast. Echter, deze resultaten zijn gebaseerd op een specifieke opzet van de experimenten, en op aannames en vereenvoudigingen welke binnen dit onderzoek gemaakt zijn. Het is mogelijk dat deze aannames en vereenvoudigingen, bijvoorbeeld de aanname dat het verkeer alleen in het horizontale vlak beweegt en het gebruik van een elementaire luchtverkeersleidersinterface, de resultaten hebben beïnvloed. Dit kan er toe geleid hebben dat de SSD grootheid wellicht te veelbelovend lijkt. Deze vereenvoudigingen waren echter nodig om: (1)

de individuele complexiteitsfactoren te isoleren en (2), eisen aan het beeldscherm (bijvoorbeeld bereik en kwaliteit van het radarbeeld), en andere eisen aan de taak (bijvoorbeeld standaardprocedures of radiocommunicatie), uit te sluiten.

Zoals eerder beschreven is het isoleren van een individuele complexiteitsfactor geen eenvoudige taak, omdat sectorcomplexiteit een ingewikkeld onderwerp is waarbij elke factor gerelateerd is aan andere factoren. Dit bemoeilijkt het systematisch onderzoeken van de effecten van afzonderlijke parameters, zonder andere parameters te wijzigen. Hoewel er getracht wordt specifieke complexiteitscomponenten te isoleren, is het mogelijk dat het onderzoek van een afzonderlijk factor (gebaseerd op een scenario van slechts twee kruisende vliegtuigen) niet leidt tot eenzelfde resultaat vergeleken met de operationele situatie. Echter, het toevoegen van extra elementen door andere, niet conflicterende, vliegtuigen in de sector, kan de aandacht van de luchtverkeersleider afleiden van het onderwerp dat wordt onderzocht. Er is dus een afweging gemaakt tussen het analyseren van afzonderlijke complexiteitsfactoren en het weergeven van een realistische verkeerssituatie aan de proefpersonen.

Daarnaast, in een poging om taken met betrekking tot het scherm en andere ongerelateerde taken te reduceren, heeft de vereenvoudiging van de instellingen van de experimenten en de functies van de simulator geleid tot simulaties die slechts een deel van het werk van de luchtverkeersleider vertegenwoordigen. Het gebrek aan realisme van de simulator kan de subjectieve werklast en strategieën van de proefpersonen beïnvloed hebben. Tevens hadden acties van de verkeersleiders geen mogelijke negatieve gevolgen, waardoor de proefpersonen over het algemeen eerder nieuwe strategieën uitprobeerden. De beperkte waarheidsgetrouwheid had ook invloed op de ervaring van gevaar en stress bij het leiden van luchtverkeer; als het niet lukte om de separatie te bewaren, dan beïnvloedde dit weliswaar hun prestaties in het experiment, maar waren er geen levens in gevaar. Vliegtuigen in de simulatie reageerden ook snel en identiek op invoer. Dit laatste kan ook invloed hebben gehad op het gedrag van verkeersleiders doordat zij instructies, om conflicten op te lossen, wellicht bewust uitstelden.

Om aan te tonen dat de methode de meest geschikte grootheid is in de bepaling van sectorcomplexiteit, dient er meer diepgaand onderzoek gedaan te worden naar de kwaliteit en robuustheid van de methode. Er zal tevens meer onderzoek gedaan moeten worden naar de complexiteit zelf, dit om meer begrip te ontwikkelen op het gebied van complexiteit en werklast. Tenslotte is het, gezien de huidige manier van opereren, noodzakelijk om het SSD met de derde dimensie (hoogte) uit te breiden.

Acknowledgements

Alhamdulillah, all praises to Allah S.W.T., The Most Greatest and The Most Merciful for the strengths and His blessings in completing my PhD journey.

Over the last six years, I have met a number of people, which have made my life colorful and interesting and also people, who have helped me go through my life as a full time student, a wife and a mother. Knowing them has made my life experience in the Netherlands to be a joyous roller-coaster ride.

First I would like to thank my promotor, Max Mulder and co-promotor, Rene van Paassen, for constantly guiding me in my work and always giving inputs and ideas for my research. My daily supervisor, Clark Borst, for the guidance, support, motivation, great discussions and contributions that are crucial to my research and thesis. I could have not made it without constant help, guidance from all of you.

My special thanks also goes to all the participants of the long hours experiments, I really appreciate your time, and effort in providing me with useful data. I know it is not easy to sit and perform your absolute best in those long hours.

My husband and love, Aswad. Who always understands how to 'cope' with me when I am all stressed out because of work. Always there as a shoulder to cry on and also knows when to push and encourage me when I need it the most. Hana, for being a lovely daughter and a blessing to our family. You know how to make my day with your apparently never-ending end-of-day storytelling.

My parents, mama and ayah, for their wholehearted love, support and never-ending dua'. My big and supportive family; Along, Kak Ina, Tiah, Tuah, Timah, Mujib, Aishah, Apad, Talha, Sarah, Dijah, Atikah, Syafiqah for always be there. Not to forget, my in-laws, mama, abah, Kak Long, Abang Wan, Abang Hafiz, Kak Liza, Kak Gee, Abang Zul, Araffat, and Nani for the prayers and encouragement for me to finish my studies.

My office mate, Maarten, Arjen and Gustavo, thanks for all the nice and long discussion about life, politics, religion, and culture. Maarten and Rolf, thanks for helping me out with the Dutch translation. My former office mate, Mark, who makes me feel welcomed at the department. To Joost, thanks for all the advice, help and interesting conversations. And not to forget, other C&S colleagues that have made coffee time something to look forward to.

Last but not least, the essential part of life in the Netherlands, my family here in the Netherlands, Pakcik Ibrahim, Kak Yah and their lovely kids, and also my female friends, who makes me feel at 'home' even though I am more than 10,000km away from 'home': Kak Ju, Kak Wan, Kak Zahirah, Kak Yan, Kak Su, Nora, Zura, Nani, Kaem, Ayu, Ija, Syu and Syud.

Without all of you in my life, the joyous roller-coaster ride will just be a ride.

Siti Mariam binti Abdul Rahman,
Delft,
February 5, 2014.

Curriculum Vitae

Siti Mariam binti Abdul Rahman was born in Leicester, United Kingdom on January 23, 1982. She was raised in Kajang, Selangor. From 1992 until 1999 she attended the SMKA Maahad Hamidiah in Kajang, obtaining her Malaysian Certificate of Education (SPM). She then received a pre-university education at Kolej Matrikulasi Melaka, Londang in 2000. In 2001, she started her bachelor degree education at Universiti Kebangsaan Malaysia (UKM), Bangi and finishes her bachelor degree (with Hons.) in 2005 in Mechanical Engineering.

In 2007, She obtained the Young Lecturer Scheme Scholarship to further her studies in Aerospace Engineering. She later decided to joined the Control & Simulation (C&S) group at Aerospace Engineering in Delft University of Technology (TU Delft). During her studies, she completed a three months internship at the National Aerospace Laboratory of the Netherlands. She graduated in November 2009.

In January 2010, she started as a PhD student in the C&S group at Aerospace Engineering in TU Delft. In the years that followed, she has work towards finishing her PhD research. She will serves at Universiti Teknologi Mara (UiTM), following her graduation as a Lecturer at the Faculty of Mechanical Engineering.