Faculty of Aerospace Engineering

Complexity Metric Comparison Study for Controller Workload Prediction in 4D Trajectory Management Environments

J. J. N. T. Toy

July 6, 2015

Delft University of Technology

Challenge the future

Complexity Metric Comparison Study for Controller Workload Prediction in 4D Trajectory Management Environments

Master of Science Thesis

For obtaining the degree of Master of Science in Aerospace Engineering
at Delft University of Technology

J. J. N. T. Toy

July 6, 2015

Faculty of Aerospace Engineering · Delft University of Technology

Delft University of Technology

Delft University Of Technology
Department Of
Control and Simulation

The undersigned hereby certify that they have read and recommend to the Faculty of Aerospace Engineering for acceptance a thesis entitled **"Complexity Metric Comparison Study for Controller Workload Prediction in 4D Trajectory Management Environments"** by **J. J. N. T. Toy** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: July 6, 2015

Readers:

____
dr.ir. M. M. van Paassen

____
dr.ir. C. Borst

____
ir. R. Klomp

____
dr.ir. A. Sharpanskykh

Acronyms

2D	two-dimensional
4D	four-dimensional
AC	aircraft count
AOL	Airspace Operations Laboratory
ARC	Ames Research Center
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATM	Air Traffic Management
DD	Dynamic Density
EID	Ecological Interface Design
FAA	Federal Aviation Administration
HOR	Horizontal Proximity Metric
ISA	Instantaneous Self Rating
LOS	loss of separation
LVNL	Luchtverkeersleiding Nederland - *Air Traffic Control the Netherlands*
MAP	Monitor Alert Parameter
NASA	National Aeronautics and Space Administration
NextGen	Next Generation Air Transportation System
NLR	Nationaal Lucht- en Ruimtevaartlaboratorium - *National Aerospace Laboratory*
PHARE	Programme for Harmonised ATM Research in EUROCONTROL
SESAR	Single European Sky ATM Research
TBO	trajectory-based operations
TBX	Trajectory-Based Complexity
TFM	traffic flow management
TU Delft	Delft University of Technology
WJHTC	William J. Hughes Technical Center

Contents

List of Figures

Thesis Outline

This document contains the Master of Science thesis of Jason J.N.T. Toy. The structure of this document contains first the IEEE journal paper in Chapter 1 where the results of the experiment study are explained in detail. Following are appendices of the research motivation in Appendix A, Travel Space tool in Appendix B, Dynamic Density metric in Appendix C, Trajectory-based Complexity metric in Appendix D, justification for the experiment study performed in Appendix E, examples of the training runs and experiment scenarios in Appendix F, the experiment documents that were used in the study in Appendix G, experiment results in Appendices H and I, and previous analysis results for the justification in Appendices J and K.

Chapter 1

IEEE Journal Paper

Complexity Metric Comparison Study for Controller Workload Prediction in 4D Trajectory Management Environments

Jason Toy, Clark Borst, Rolf Klomp, Max Mulder, and René van Paassen

***Abstract*—The future of air traffic management moving to 4D trajectory-based operations will require the development of new airspace sectors to increase aircraft capacity and advanced 'human-centered' decision support tools for future air traffic controllers. To evaluate and aid in the design of future air traffic management systems, complexity metrics would help to speed up the development of advanced safe air traffic management systems. Although an airspace sector may look 'complex' with many aircraft, it does not equivocate to actually being complex with the right tool. The complexity of a sector or traffic scenarios depends on a large amount of factors, irrelevant of the tool. Previous studies with well developed state-based complexity metrics focused on air traffic controllers safely controlling traffic of today, with a 'hands-on' approach. One recent metric based on trajectory-based management could prove to help predict controller workload. The goal of this study has been to empirically investigate if a complexity metric can predict human controller workload in future 4D trajectory management environments. For this purpose a previously developed 4D management tool had been used to support a controller in an envisioned future large airspace sector with varying traffic structures and perturbation levels. A well developed state-based complexity metric was compared against a recent trajectory-based complexity metric by the results of the reported workload experienced. Results of a human-in-the-loop experiment despite required time-shift indicate that the trajectory-based complexity metric looks promising to workload predictions in 4D trajectory management environments.**

***Index Terms*—workload, inherent complexity, 4D trajectories, air traffic management, human-machine interaction**

I. Introduction

THE current evolution of the Air Traffic Management (ATM) system is foreseen to bring a paradigm shift to the work domain of the Air Traffic Controller (ATCO) [1], [2]. With the addition of *time*, future high-precision four-dimensional (4D) trajectories Air Traffic Control (ATC) tools are being developed throughout the world for the ATCO to be encompassed within a Trajectory-Based Operations (TBO) environment in ATM. These 4D trajectory tools will allow ATCO to plan aircraft movements farther in advance and make the behavior of the system more predictable. This will allow ATM to cope with expected increasing traffic volume and a switch to strategic management for ATCO rather than the current 'hands-on' method of control.

It is believed that by improving measures of ATC complexity it can benefit the evaluation of ATM productivity, benchmarking cost effectiveness, assessment of the impact of new tools and procedures, and airspace redesign. Numerous individual components greatly impact the elaborate connection among complexity and workload. Although there are difficulties for fully capturing the notion of cognitive complexity mathematically it does not mean that the human factor must remain unknown [3]–[11]. In ATM, complexity is made up of many factors that contribute to the level of complexity for an ATCO. There are two types of complexity; inherent and apparent. *Inherent* complexity are the intricate qualities that arise from the airspace properties such as weather, terrain, airspace restrictions, traffic density, traffic flows, aircraft performance characteristics, abnormal events, etc. *Apparent* complexity are the intricate qualities that come from the interface to the controller such as mono-color and multi-color displays, touch screens, physical arrangements of displays or consoles, control room layout, software used to display information, etc. The focus of this research was on the inherent complexity as there is no clear winner of the future ATM system interface to divulge and investigate its apparent complexity.

Evaluating which of these 'human-centric' 4D trajectory-based tools (currently being developed worldwide) that will be best suited for a particular airspace sector(s) is difficult to determine as there are no proven metrics to help evaluate a sector's inherent complexity in TBO environments. This is new territory with a lot of unknowns and having a complexity metric that could predict what levels of workload an ATCO is experiencing is crucial in the design of new sectors and the development of 4D trajectory-based tools. There has been extensive research into complexity metrics developed with state-based approaches for the current ATM operations. However, research into this new strategic management heavily aided with automation for a TBO environment is still in its early stages.

In this paper we investigate a trajectory-based complexity metric called Trajectory-Based Complexity (TBX) [12] against a state-based complexity metric called Dynamic Density (DD) [13] to see how well they compare with workload ratings, from a user group performing as ATCOs in a TBO environment using a Travel Space Representation 4D trajectory-based tool [14]. There are no expert ATCOs yet in 4D ATM so it is unknown how new air sectors and ATM tools will perform. The goal of this paper is to empirically investigate which equations of airspace complexity could predict air traffic controller workload in 4D trajectory management environments in scenario runs with varying scales of perturbations and traffic orderliness.

J. Toy, C. Borst, R. Klomp, M. Mulder, and M.M. van Paassen are with the Control and Simulation Section, Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS Delft, The Netherlands e-mail: J.J.Toy@student.tudelft.nl.

The structure of the paper is as follows. First, the explanation of the practical use of the Travel Space tool used by human controllers will be discussed. Next, the form of the DD metric used for the TBO study will be elaborated on followed by the TBX metric. Then the experimental design is presented along with results, discussion, and conclusions.

II. Travel Space Representation

The Delft University of Technology (TU Delft) Travel Space Representation is a Direct Manipulation Interface (DMI) 4D trajectory management tool that visualizes the planned trajectory path of an aircraft within the sector, as well as visually displaying all possible airspaces of safe control (based on aircraft performance profile) for an aircraft that an ATCO can select and modify to maneuver in the future [14].

The Travel Space tool assists the ATCO by showing the shared representation of the solution space of a given aircraft to resolve conflicts with a certain look ahead time as seen in Figure 1. The solution space has shapes of ellipses that grow outward with increasing velocity. The size and shape of the solution space are determined by aircraft performance constraints which represents the space in which the selected aircraft can be rerouted without exceeding its speed envelope or bank angle limits as seen in Figure 1b. The areas of the solution space are colored green for rerouting solutions to conflicts or red for conflict zones (resulting from other traffic) which must be avoided as seen in Figure 1b, where the light gray area represents green and the dark gray area represents red. It is still up to the human controller to select anywhere in the solution space and confirm a trajectory by means of a mouse input and keyboard key combinations input devices. A placement of a new waypoint is tentatively selected to ensure separation. Once confirmed this creates a new waypoint that will create two new solution space segments as shown in Figure 1c. All aircraft within the sector show their complete trajectory path in gray to their exit point throughout the whole runtime, giving constant situational awareness for controllers. The restricted areas are outlined in red but do not explicitly show up as a conflict in the Travel Space solution airspace when selecting an aircraft, the ATCO actively needs to see if any trajectory paths crosses restricted areas. A more detailed description of the Travel Space Representation can be found in [14]. This tool will be used to manage traffic in the study.

III. Dynamic Density Metric

DD tries to capture the complexity or difficulty of a traffic situation and is a collective effect of all factors, or variables, that contribute to the sector level ATC complexity at any given time. It can also be described as a ATC taskload which is a function of the number of aircraft and the complexity of traffic patterns in a volume of airspace [13], [15]. This metric is widely used to predict current ATC/sector complexity. However, this metric is state-based and not originally intended for 4D trajectory operations. When using DD a static snapshot of a given traffic scenario can be analyzed, but the dynamics of the system (given time-based intent) are limited (via state-extrapolation parameters) which trajectory management would highly make use of.

TABLE I
Valid NASA DD Metric 1 Elements for Travel Space

Element	Renamed	Description
C_1	Flow	Number of Aircraft Flow
C_5	HOR1	Inverse Mean Weighted Horizontal Separation Distance
C_7	HOR2	Inverse of the Average Minimum Horizontal Separation Distance
C_9	HOR3	Inverse of the Minimum Horizontal Separation
C_{11}	TTG1	Time-To-Go Fraction of Aircraft
C_{12}	TTG2	Time-To-Go Inverse of Average Minimum
C_{13}	TTG3	Time-To-Go Inverse of Smallest
C_{14}	SPD1	Variance of Groundspeed
C_{15}	SPD2	Groundspeed Ratio of Standard Deviation to Mean
C_{16}	MWN	Mean Conflict Resolution Difficulty

There have been multiple organizational research efforts that developed and validated several flavors of DD metrics. With the previous experiments [16] using DD it had been shown that National Aeronautics and Space Administration (NASA) Metric 1 was the most promising of the DD metrics for the experimental conditions used for the Travel Space tool. Only the NASA Metric 1 had been considered in the experimental study and is explained in more detail in the following sections. All DD metrics are more or less similar in the basis, only some metrics use more or other parameters.

A. NASA Metric 1

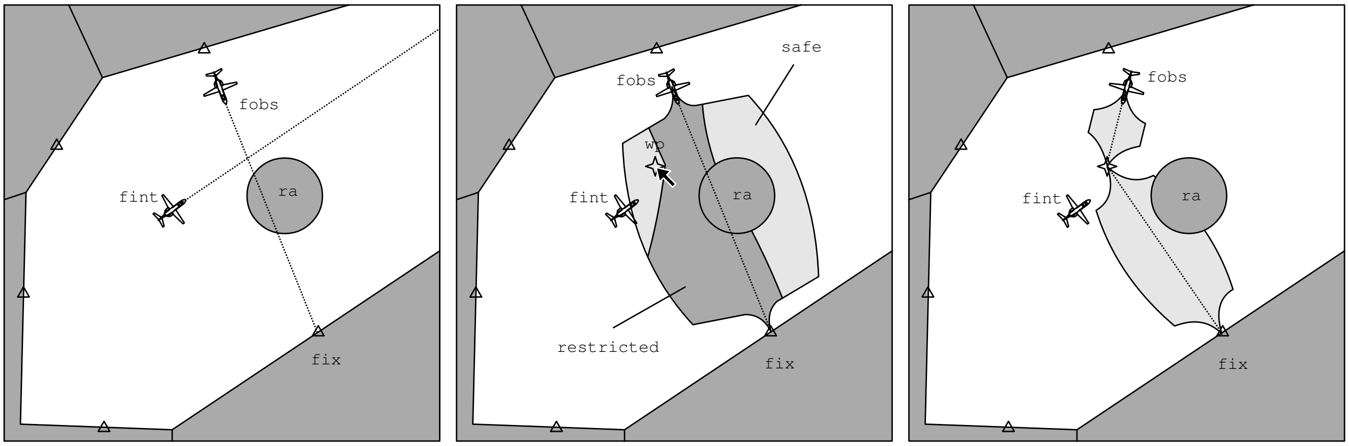
The original metric consisted of sixteen complexity elements that are summed to give an overall value of complexity [15]. However, for the scope of this study it is assumed that all aircraft are at the same flight level with no possibility of vertical manipulations. Therefore, this reduced the valid elements to ten for this study. An overview of these parameters are given in Table I. The modified DD NASA Metric 1 equation for the Travel Space tool is as follows:

$$\begin{aligned} DD_{TS} &= a*Flow + b*HOR1 + c*HOR2 + d*HOR3 \\ &+ e*TTG1 + f*TTG2 + g*TTG3 \\ &+ h*SPD1 + i*SPD2 + j*MWN \end{aligned} \tag{1}$$

Where the weighting coefficients (a, b, c,...) are determined afterwards using linear regressions to fit the DD to a specific sector and are valid only for that sector. $Flow = N/N_{max}$, where N is the total number of aircraft within the sector at any instant of time and N_{max} is the acceptable maximum number of aircraft in the sector. The following sections go into detail about the modified elements of (1).

B. Horizontal Proximity Metric 1 (HOR1)

The inverse of the mean weighted horizontal separation between aircraft pairs is one of several proximity measures. The reasoning for using the inverse is that decreasing mean distances results from reduced separation between neighboring



(a) An en-route traffic scenario with two conflicting aircraft and a restricted airspace.

(b) The TSR and placement of an intermediate waypoint

(c) Resulting trajectory for the observed aircraft

Fig. 1. Travel Space Representation (TSR) support for the task of manual trajectory revision of an observed aircraft by air traffic control

aircraft. Following is the reduced equation, taking into account that all aircraft are at the same flight level for the study.

$$HOR1 = \frac{N}{\sum_{1 \leq i \leq N} \left(\frac{\sum_{1 \leq j \leq N} d_{ij}^{-1}}{\sum_{1 \leq j \leq N} d_{ij}^{-2}} \right)} \tag{2}$$

Where N is the number of aircraft within the sector, d_{ij} is the horizontal separation distance between the two aircraft i and j. As the mean weighted separation distances decrease this will cause the value of $HOR1$ to increase, adding to the complexity contribution to the complete DD equation.

C. Horizontal Proximity Measure 2 (HOR2)

The inverse average minimum horizontal separation between aircraft pairs is the second horizontal proximity measure. This measure is based on the average minimum separation between aircraft that would give more workload to the controller if on average the minimum separation are decreasing. The reduced equation is as follows:

$$HOR2 = \frac{N}{\sum_{1 \leq i \leq N} \min\{d_{ij}\}} \tag{3}$$

Where N is the number of aircraft within the sector, d_{ij} is the lateral distance between the i and j aircraft. The numerator counts the number of aircraft for which at least one other aircraft is found within its altitude neighborhood. As the average minimum distances decrease, the value of $HOR2$ will increase, adding to the complexity contribution of the complete DD equation.

D. Horizontal Proximity Measure (HOR3)

The inverse of minimum horizontal separation in the same vertical neighborhood is the final horizontal proximity measure. This measure is based on the minimum separation for a pair of aircraft within the group. This close separation between a pair of aircraft would cause the controller to focus the attention on this pair of aircraft because of the possibility of separation violation. The reduced equation is as follows:

$$HOR3 = \frac{1}{\min_{1 \leq i \leq N} \{d_{ij}\}} \tag{4}$$

Where N is the number of aircraft within the sector and d_{ij} is the lateral distance between aircraft i and aircraft j. As the smallest distance between an aircraft pair out of the entire group decreases, the value of $HOR3$ will increase, contributing more to the complete DD equation.

E. Time-To-Go to Conflict Measure 1 (TTG1)

The fraction of aircraft with time-to-go to conflict less than $\Delta t = 600s$ is the first of these kinds of measures that consider the urgency of conflict resolution. For this state-based approach the number and frequency of the sector controller's actions for conflict resolution can be based on the relative position, heading, and speed of aircraft pairs. These three qualities together determine if the aircraft pair would violate the separation minimum in the future.

The rate of change of distance between a pair of aircraft can be used to determine if the pair is moving toward each other or away from each other. In the cases where the aircraft are moving toward each other the pairs that are in close proximity are more important compared to those aircraft pairs that are farther away. The closer pairs require immediate conflict resolution and the further away pairs can be postponed. Following is the reduced equation for the range rate:

$$\dot{d}_{ijTS} = \frac{(d_{xij} V_{xij} + d_{yij} V_{yij})}{d_{ij}} \tag{5}$$

Where d_{ij} is the distance between the i and j aircraft pair and the terms d_{xij} and d_{yij} are the distance coordinates and V_{xij} and V_{yij} are the relative velocity components with respect to the body frame, attached to the i aircraft, measured along the inertial frame axes. The time-to-go to conflict t_{ij} can be determined in terms of the range rate given by (5) as:

$$t_{ij} = -\frac{d_{ij}}{\dot{d}_{ijTS}} \tag{6}$$

From (6), it can be seen that time-to-go is positive if the closing rate is negative which indicates that the pair of aircraft are converging. From this equation several measures can be developed. The threshold time when conflict resolution becomes urgent is defined as Δt and T_i is the set of neighboring aircraft with time-to-go less than or equal to Δt seen as follows:

$$T_i = \{j | 0 \leq t_{ij} \leq \Delta t;\ \ j \neq i\} \tag{7}$$

Increasing numbers of such pairs can result in the monitoring workload to increase. In other words, the ATCO has to attend to urgent problems which will increase workload. Following is the reduced flow complexity measure, $TTG1$, derived from this increase. $TTG1$ is based on the number of positive time-to-go less than or equal to Δt.

$$TTG1 = \frac{\sum\limits_{1 \leq i \leq N} \sum\limits_{j \in T_i} 1}{2N} \tag{8}$$

N is the number of aircraft within the sector and the factor of 2 is used in the denominator because the pairs i an j are counted twice. The term T_i is the indices with time-to-go less than the threshold value, refer to (7). The time-to-go threshold Δt is set to 600s, t_{ij} is the time-to-go, d_{ij} is the distance between the i and j aircraft pair, $\dot{d}_{ijTS}$ is the range rate. Notice as the number of aircraft pairs increase the value of $TTG1$ increases, adding to the complexity of DD.

F. Time-To-Go to Conflict Measure 2 (TTG2)

The inverse minimum time-to-go to conflict with less than $\Delta t = 600s$ is another measure that considers the urgency of conflict resolution. This measurement is based on the average time-to-go value on a given set of aircraft pairs with less than 600 seconds which indicates the time the controller has for resolving conflict in general. Following is the reduced equation:

$$TTG2 = \frac{\sum\limits_{1 \leq i \leq N} [j \in T_i]}{\sum\limits_{1 \leq i \leq N} \min\limits_{j \in T_i} \{t_{ij}\}} \tag{9}$$

N is the number of aircraft within the sector. The term T_i is the set of neighboring aircraft with time-to-go less than or equal to 600 seconds, refer to (7). The time-to-go to conflict is t_{ij}, refer to (5) and (6). The numerator counts the number of aircraft that have at least one other aircraft in the set defined by (7). The index of summation i is bound by 1 and N therefore, the maximum value of the numerator is N. For the denominator the minimum time-to-go between the i aircraft and all members in the set T_i is computed for every aircraft. The resulting values are summed up to obtain the value for the denominator. If there are no conflicts within 600 seconds, the value of the measure is zero. When the summed minimum times to conflict between aircraft pairs become gradually smaller, this measure increases.

G. Time-To-Go to Conflict Measure 3 (TTG3)

The inverse of smallest time-to-go conflict for aircraft pairs with time-to-go to conflict less than $\Delta t = 600s$ is similar to (4). This measurement is based on the logic that the workload experienced by the controller will be higher, if the time available for resolution of the immediate conflict is smaller. Following is the reduced equation:

$$TTG3 = \frac{1}{\min\limits_{1 \leq i \leq N} \left\{ \min\limits_{j \in T_i} \{t_{ij}\} \right\}} \tag{10}$$

N is the number of aircraft within the sector. The term T_i is the set of neighboring aircraft with time-to-go less than or equal to 600s, refer to (7). The time-to-go is t_{ij}, refer to (5) and (6). If there are no conflicts within 600 seconds, the value of the measure is zero. When the average time to conflict between aircraft pairs becomes smaller (i.e., conflict is more imminent), this measure increases the overall DD score.

H. Speed Measure 1 (SPD1)

The variance of groundspeed is how far the values of aircraft groundspeed are spread out amongst each other. For DD it is thought that the variance of groundspeed is possibly a good measure of complexity. Average speed and average sector transit time have been used as measures of complexity by some researchers in the past but a study reviewing these measures for DD concluded that these may not be a good measure of complexity because the variability in the parameters within the sector was much lower than between sectors [15]. The study also found that the mean airspeed was not significantly correlated to the behavioral response so it did not give a good measure of job difficulty. The variance of groundspeed equation reduced without vertical components is as follows:

$$SPD1 = \sigma_{vg}^2 = \frac{\sum\limits_{1 \leq i \leq N} (V_i - \bar{V})^2}{(N-1)} \tag{11}$$

$$V_i = \sqrt{V_{xi}^2 + V_{yi}^2} \tag{12}$$

$$\bar{V} = \frac{\sum\limits_{1 \leq i \leq N} V_i}{N} \tag{13}$$

In (11), σ_{vg}^2 is the variance of groundspeed. N is the number of aircraft within the sector. V_i is the groundspeed of aircraft i and $\bar{V}$ is the mean of groundspeed. The groundspeed and mean groundspeed are defined respectively in (12) and (13). In the groundspeed equation velocity components are given in terms of x and y for an i^{th} aircraft. The value can have a wide range depending on the amount of aircraft in the sector, affecting the average speed value and the speed differences. This measure contributes to the complete DD equation and can have a large value although each DD measure has a coefficient fitted to the sector which is determined after numerous runs and the data linearly regressed to workload.

I. Speed Measure 2 (SPD2)

The ratio of standard deviation of ground speed to mean of ground speed is known as the contrast ratio. The measure models the heuristic that higher variance does not increase workload if the average groundspeed is low. The measure is based on the variance and the mean of the groundspeed which is as follows:

$$\sigma_{vg} = \sqrt{\frac{\sum\limits_{1 \le i \le N} (V_i - \bar{V})^2}{(N-1)}} \tag{14}$$

$$SPD2 = \frac{\sigma_{vg}}{\bar{V}} \tag{15}$$

In (14), N is the number of aircraft within the sector and V_i is the groundspeed of the i^{th} aircraft. In (15), σ_{vg} is the standard deviation and $\bar{V}$ is the mean groundspeed, (13). The value of $SPD2$ becomes very small, a fraction of the value one, because of the square root and divided by the mean groundspeed. This measure contributes to the complete DD equation and although the value is small, each DD measure has a coefficient fitted to the sector which is determined after numerous runs and the data linearly regressed to workload.

J. Mean Conflict Resolution Difficulty (MWN)

The mean conflict resolution difficulty is based on crossing angle of two aircraft. Small crossing angle conflicts are the most complex and 90 degree crossing angle conflicts are the least complex. With head-on conflicts to be high complexity because of high closing rates. Shallow crossing angle conflicts are harder to detect but easier to resolve because the controller has more time for path corrections while head-on conflicts are easy to detect but harder to resolve. A study has found shallow crossing angle resolutions have to be initiated earlier compared to large crossing angles [15]. A normalization factor, n_f, of $3208.2s$ (53.47 min) has been used based on a graph of normalized time of resolution initiation as a function of cross angle [15].

The crossing angle ξ_{ij} for the i and j pair of aircraft is given by $\xi_{ij} = min(|\chi_{ij}|, 2\pi - |\chi_{ij}|)$ where χ_{ij} is the relative heading angle of the converging pair of aircraft. The time-to-go t_{ij} as defined in (6) and using (5) is less than the time-to-go threshold Δt in the set T_i defined in (7). The heading angle for i aircraft is defined below.

$$\chi_i = tan^{-1}\left(\frac{V_{yi}}{V_{xi}}\right) \tag{16}$$

Where V_{xi} and V_{yi} are the velocity vector components on the horizontal plane. The level of resolution difficulty as a function of crossing angle is obtained from the mentioned graph which is available as a lookup table. Each of the converging pair of aircraft can be combined to define the overall complexity measure associated with conflict resolution. The modified mean conflict resolution difficulty is as follows:

$$MWN = \frac{\sum\limits_{1 \le i \le N} \sum\limits_{j \in T_i} \varpi_{\xi_{ij}} n_f}{2N} \tag{17}$$

N is the number of aircraft in the sector and T_i is the set of neighboring aircraft with time-to-go less than or equal to $\Delta t = 900s$. The level of resolution difficulty is given as $\varpi_{\xi_{ij}}$ and the normalization factor is $n_f = 3208.2s$. A factor of 2 has been included in the denominator to account for the fact that i and j aircraft pair is counted twice. When there are head-on or small crossing angle conflicts between a pair or pairs of aircraft, this measure increases the overall DD score.

IV. Trajectory-based Complexity Metric

TBX is a modified aircraft count and a new approach to measure traffic complexity that unlike DD can be computed and communicated easily to predict in real-time sector complexity for a TBO environment [12]. Previous metrics of traffic complexity such as DD are state-based. TBX is a recent metric developed and evaluated in the Airspace Operations Laboratory (AOL) at NASA Ames Research Center (ARC) during human-in-the-loop studies of trajectory-based concepts since 2009. This metric considers more dynamic factors such as weather, aircraft equipage, predicted separation violations, as well as static factors such as sector size. It is said to be a better predictor of workload than aircraft count and works well for complexity management in trajectory-based operations that can readily adjust to future operational concepts [12]. Other metrics of complexity in the past such as DD had to be computed off-line and matched to previously collected data rather than in real-time based on current and predicted data. Resulting metrics that provided good statistical fit to the historical data were often ill-suited to characterize the uncertainties associated with traffic predictions [12].

The TBX metrics were design to: be comprehensible to human operators, have weightings that would be easy to modify for new equipage and operational environments, be stable over the prediction time horizon, use TBO to create more stable trajectory predictions over a 1-2 hour time horizon, and combine different complexity factors into a single modified aircraft count value that human operators could use in conjunction with a Monitor Alert Parameter (MAP) value. To support real-time traffic flow management (TFM), TBX was designed for real-time computations of complexity estimates based upon predictions of 4D trajectories for weather and aircraft.

The drive for this comparison study of metrics were due to these promising claims of the TBX metric for TBO environments. Before the TBX can be computed it requires prior determinations of defining the nominal conditions, where aircraft count is a good predictor of controller workload, and defining adjustments to the nominal conditions that capture the difference in complexity between the current conditions and the nominal conditions. An airspace sector's predicted aircraft count is used to determine the maximum traffic threshold that a controller can handle in that sector; this maximum threshold value is called MAP and is set based on average sector flight time.

The adjustments are ratios that describe the relationship between the current value and the nominal value. There are two types of adjustments, primary complexity adjustments and secondary complexity adjustments. Primary complexity adjustments are factors that can cause a similar effect to the complexity as the aircraft count, can offset the aircraft count, or have a detrimental impact. Theses factors are expected to have a multiplicative effect on the aircraft count. Weather or sector size are an example of dominant items that have primary adjustments. Secondary adjustments are factors that have a secondary/smaller impact on complexity which can be specific to a given operational environment. Their impact is less dependent on the overall aircraft count such as number of conflicts or transitioning aircraft [12]. The TBX value of a sector *s* at time *t* equals the sum of its predicted aircraft count (*ac*) multiplied by the product of the primary adjustments (*px*) and the nominal aircraft count multiplied by the weighted sum of the secondary adjustments (*sx*) minus one. Following is the general form of TBX:

$$tbx(s,t) = ac(s,t) * \prod_{i=0}^{n} px_i(s,t) + ac_{nom}(s) * \left(\frac{\sum_{j=0}^{m}(w_j * sx_j(s,t))}{\sum_{j=0}^{m}(w_j)} - 1 \right) \quad (18)$$

$$tbx_{nom}(s,t) = ac(s,t) \quad (19)$$

Where $ac(s,t)$ is the predicted aircraft count of sector s at time t, $px_i(s,t)$ is the adjustment for primary complexity item i in sector s at time t, $ac_{nom}(s)$ is the nominal aircraft count in sector s, w_j is the weight of secondary complexity item j, and $sx_j(s,t)$ is the adjustment for secondary complexity item j in sector s at time t. TBX and its adjustments have the value of 1 under nominal conditions which satisfies (19). All adjustments px_i and sx_j are limited to a range of 0 to 2 to prevent one factor from dominating the overall TBX value [12]. In this structure of (18), MAP values can be based on nominal conditions for a given sector and then compare aircraft count and TBX value relative to the MAP to assess the peak that exceeds the MAP. This representation has the same format as aircraft count making the interpretation of TBX simple.

V. TBX CALCULATION

For this study, (18) has the complexity factors of aircraft count, sector area, aircraft predicted to penetrate weather, and predicted loss of separation (LOS) events. Following is the modified equation for this study:

$$tbx(s,t)_{TravSP} = ac(s,t) * px_{sa}(s,t) * px_{wx}(s,t) + ac_{nom}(s) * \left(\frac{sx_{PrLos}(s,t)}{1} - 1 \right) \quad (20)$$

Where $ac(s,t) = N$, which is the number of predicted aircraft count of sector s at time t and is the dominant complexity factor. The term $ac_{nom}(s)$ is the nominal aircraft count ($ac_{nom}(s) = 36$) in sector s defined to have 80% of the MAP value (MAP = 45) throughout the complexity factors. The following complexity factor sections explain the rest of the terms in (20) in detail.

A. Sector Area (px_{sa})

The usable sector area can have an influence on complexity since larger usable sectors provide more airspace than smaller ones for maneuvering aircrafts, communication, and time to decide on new/modified trajectories [12]. Following is the sector area element of TBX:

$$px_{sa}(s,t) = \left(\frac{sa(s,t)}{sa(s,t) - sua_{rs}(s,t)} \right)^{0.15} \quad (21)$$

Where $px_{sa}(s,t)$ is a primary complexity adjustment for sector area in sector s at time t. The term $sa(s,t)$ is the sector area of the current sector s at time t. The term $sua_{rs}(s,t)$ is the special use airspace area (e.g., weather, government zones, military zones, etc.) or restricted area for the current sector s at time t. In this study only static restricted areas will be considered. If there were multiple kinds of special use airspace areas then these would be added with each other (e.g., $sua_{wx}(s,t) + sua_{gov}(s,t) + sua_{mil}(s,t)$). The power of 0.15 has been chosen so that the $px_{sa}(s,t)$ value does not dominate the overall tbx_{TravSP} value. The power value can be tuned empirically with more studies by assessing the impact of sector size on the overall complexity value and comparing the complexity to the corresponding controller workload. Notice if there are no restricted areas, $sua_{rs}(s,t) = 0$, then the value of px_{sa} becomes 1 and does not contribute to the tbx_{TravSP} value. As the value of $sua_{rs}(s,t)$ increases the px_{sa} value rises exponentially. In other words, the bigger the restricted areas become the more complex it will be to reroute air traffic with the decreasing usable sector area.

B. Weather (px_{wx})

Hazardous weather impacts complexity by restricting the available airspace, which causes the need for rerouting flights and often results in substantial increases in communication between pilots and controllers. Weather becomes less predictable the further out into the future the prediction is made [12]. The number of aircraft predicted to penetrate weather within a selected sector is used in the TBX calculation. The predicted weather trajectory is compared against the predicted aircraft trajectory to determine whether these two intersect within a sector. Following is the weather element of TBX:

$$px_{wx}(s,t) = 1 + 2 * \left(\frac{ac_{wx}(s,t)}{MAP} \right) \quad (22)$$

Where $px_{wx}(s,t)$ is a primary adjustment for the weather penetrations in sector s and at time t. The term $ac_{wx}(s,t)$ is the number of aircraft predicted to penetrate the weather of the current sector s at time t. The MAP value is set to 45 aircraft. Notice if half of a sectors MAP value will penetrate the weather, the complexity is doubled in comparison. If the maximum value of aircraft were to penetrate the weather, $ac_{wx}(s,t) = MAP$, it would triple the complexity.

C. Conflicts (sx_{PrLos})

When aircraft become in conflict it adds to complexity by requiring the controller to assess the situation and resolve the conflict by maneuvering at least one of the conflicting aircraft. Usually this includes coordinating with adjacent sectors which further increase the workload. Trajectory calculations are used to predict conflicts only reliably for up to 20 or 30 minutes because of uncertainties [12]. With the data collection setup for the experiment only one type of conflict related data was considered because the tool only focuses on conflicts predicted within the sector and does not account for aircraft that have a conflict in a different sector. Following is the conflicts element of TBX for the number of predicted LOS events:

$$sx_{PrLos}(s,t) = 1 + 2 * \left(\frac{PrLos(s,t)}{MAP} - PrLos_{nom} \right) \quad (23)$$

Where $sx_{PrLos}(s,t)$ is a secondary adjustment for the predicted number of LOS that occur in sector s at time t. The term $PrLos(s,t)$ is the number of predicted LOS events to occur in the sector s at time t, if no action is taken. The term $PrLos_{nom}$ is the nominal value for predicted LOS to occur in a sector s at time t. The value of $PrLos_{nom}$ is set to 0.05 for the Travel Space experiment, which means the number of conflicts is about as much as 5% of the MAP value (45 aircraft). Notice the limit of $sx_{PrLos} = 2.9$ when $PrLos = MAP$ and in the secondary adjustment portion of (20) it reduces to 1.9.

VI. Human-in-the-Loop Experiment

To evaluate which, if any, of the two complexity metrics could predict human controller workload in the future of ATC, 4D trajectory management environment, a human-in-the-loop experiment was performed. The Travel Space tool was used to serve as the future 4D trajectory tool for this experiment. Participants were asked to manage various scenarios of trajectory-based air traffic without the aid of any automated advisories (i.e., by using the Travel Space tool alone). The goal of the experiment was to investigate if the state-based DD metric and/or the trajectory-based TBX metric could predict human controller workload in 4D trajectory management under varying traffic and perturbation levels within the sector and whether it is more suitable in some cases than for others.

A. Participants

This experiment assumed that future ATCOs will be controlling quite differently using strategic management rather than what controllers of today currently use with a proactive 'hands-on' approach. So as to minimizing skilled ATCOs dealing with current sized airspace sectors and aircraft loads (which both are expected to greatly increase) and their tendencies to control with a 'hands-on' approach, participants that have no formal ATCO experience were used in this study. This was to help ensure that the data was not affected from old habits of structuring or rerouting aircraft as ATCOs of today are trained to do, possibly disregarding the advantages of the 4D trajectory management environment. As a first screening, potential participants were first asked if they have color normal vision as a color deficient person (depending on the severity) may have trouble discriminating the red and green visuals (resulting in a slower reaction time) that predominantly appear in the Travel Space tool. The experiment was performed with a total of sixteen novice participants (12 males, 4 females, average age of 25) who are all aerospace master students with some knowledge of how ATM functions but no prior training in operational air traffic control.

B. Procedure

Participants were first given an initial briefing of the importance of 4D trajectory management in the future and to this experiment followed with an outline of the experiment. They were asked to fill out a questionnaire form about their brief background and a consent form of their participation. Next a 30-45 minute training session was given during which the participants were asked to follow an interactive script facilitated by the researcher to become familiarized with the Travel Space tool and its functionality. The training ended when the script was completed, and when the participants indicated that they had a good understanding of how to use the Travel Space tool to manage airspace traffic. This was determined by their performance of short example scenarios to see how comfortable and confident they were using the tool and repeated if necessary. Once training was completed it was followed with a short break of approximately 10 minutes.

In the main experiment participants were asked to manage airspace traffic within a fictional two-dimensional sector under various initial conditions. There were four different scenario conditions in total. For each run the goal was to reroute traffic safely (i.e., without losses of separation or restricted area intrusions) within the controlled sector. Participants were informed that optimal routes or the amount of rerouting waypoints was not a criteria, as long as aircraft met their target exit point safely. After the initialization of a scenario, participants controlled airspace traffic through the use of issuing planned route changes to the 4D trajectories of each individual aircraft by manipulating waypoints using the Travel Space tool. The resulting trajectories were automatically calculated and executed by the aircraft upon the confirmation from the participant. Throughout each experiment run, every 25 seconds the participant was asked to rate their current workload rating from an Instantaneous Self Rating (ISA) Scale that popped up with an audible tone in a side bar on the screen scaled from 0 to 100. Although this might have been contributing to workload/stress to the participant, it was necessary to collect readings at these intervals to get a good resolution in order to compare the calculated metrics with workload data.

Halfway in between the four experiment runs a short break of approximately 10 minutes was given. At the completion of the last run a debriefing allowed the participant to provide feedback of the scenario runs and their control strategy.

C. Apparatus

The experiment was performed at the TU Delft Aerospace Engineering Control and Simulation department in the ATM lab on a dedicated software-based ATM platform, running on a single computer. The Travel Space tool was integrated in a traditional plan-view display (PVD) that provided a top-down view of the airspace sector and air traffic. The Travel Space tool was presented on a 30 inch screen (60 Hz LED, 2560 x 1600 pixels) placed in front of the participant. The input was given by a standard mouse and keyboard input devices with mouse and key combinations giving control options.

D. Independent Variables

The experiment had two within-subject independent variables that where:

- *Orderliness:* The initial air traffic orderliness, with two levels: structured traffic and unstructured traffic, and
- *Perturbation:* The number of aircraft in the air traffic sample that were required to be rerouted in order to prevent losses of separation or restricted airspace intrusions, with two levels: small perturbation and large perturbation.

The rationale for the *orderliness* variable is due to the form in which TBO in general will be implemented in the future is not yet defined. Therefore, both structured (fixed route-structures) and unstructured (free-routing) traffic conditions have been considered in this research which could help to see in which scenario the metrics correlate best with workload. In structured air traffic, all aircraft traversed the sector through two main highway bidirectional predictable streams that crossed each other approximately perpendicularly around the middle of the sector. With Travel Space tool experiments in the past the structured air traffic condition had several traffic streams going to various points and angles which is not tightly strict as for the structured case in this experiment. To throughly test the metrics the most tightly structured case was created with a balance of traffic going both directions and cases of LOS if the human controller does not put any inputs to actively reroute traffic, hence the two highway streams crossing perpendicularly. In unstructured traffic, the same amount of aircraft entered the sector at the exact same times as the structured case but uniquely arranged so that each aircraft entered various entry points in the sector at different headings to appear very unstructured. The scale of the sector was also vastly larger compared to previous experiments to increase the entry/exit point options for aircraft making it possible for an extremely unstructured case and also to make sure that there was a wide spread in workload. The result was participants heavily relying on the Travel Space tool to resolve conflicts and being situationally aware of all aircraft headings at anytime. Only one baseline structured and one baseline unstructured initial traffic scenario were used in all four scenarios.

The *perturbation* variable was defined by the minimum number of aircraft that the participant initially had to re-align in order to resolve all conflicts and restricted area violations. The initial conflicts were purely geometrical; all aircraft entered the sector with the same speed, thus catch-up and/or overtake scenarios were not considered. In the small perturbation condition with structured traffic (S_S, where the subscript is the perturbation), a restricted area (circular area with a 25 NM radius) was added in the middle of the sector and in the path of two of the four one-way streams that are perpendicular to each other. This required active rerouting and depending on the participant's rerouting waypoint placements it could cause more LOS later in the future. In the small perturbation condition with the unstructured air traffic (U_S), aircraft frequently had initial headings to cross the restricted area and also depending on the participants rerouting choices aircraft could cause more LOS in the future. For structured and unstructured traffic both had approximately the same number of aircraft to reroute. The large perturbation condition with structured traffic (S_L) had three restricted areas that totaled three times as big as the small perturbation scenario. The restricted areas which were circular areas were placed in a triangular formation in the middle of the sector with sufficient spacing in between each other to allow rerouting solutions in between the restricted areas for the participant. The initial headings of the structured traffic streams crossed the restricted areas. For the final scenario of large perturbation and unstructured traffic (U_L), the restricted areas were in exactly the same placement as described before but with aircraft entering the same exact times but at randomized entry points. The initial aircraft headings when entering the sector were frequently crossing the restricted areas and amount of LOS depended on the participants rerouting choices.

The rationale for no middle perturbation condition was due to requiring the experiment time to be extended which could discouraging willing participants for the study, their attention span performance degrading, and fighting for the availability of scheduling lab time between other research experiments to be performed in the lab. Another consideration was the scenario times being shortened but because of the large size of the sector there would not be ample time for aircraft build up and LOS to occur due to participants rerouting choices. However, the most important of all was the timing intervals of the ISA rating measurements would either need to be shortened for better resolution that would cause more of a burden to participants to just complete the primary task or have less readings to compare the metrics to which would not be a good enough resolution to draw much conclusions from.

The control variables in the experiment were: the sector area, size and shape, the availability of the Travel Space tool, the size and shape of the restricted area, the initial traffic sample in structured and unstructured conditions, and all aircraft entered the sector the exact same times for all scenarios.

E. Dependent Measures

The following dependent measure were used to investigate the effect the traffic orderliness and the perturbation level in a 4D trajectory management environment on the human-controller workload:

- *Trajectory-based Complexity (TBX):* Calculation of the complexity in the sector at an instance of time based on a modified aircraft count using a 4D trajectory approach,
- *Dynamic Density (DD):* Calculation of the complexity in the sector at an instance of time using a state-based approach,

TABLE II
DEFINITION OF THE FOUR EXPERIMENT CONDITIONS.

Condition	Orderliness	Perturbation
S_L	structured	large
S_S	structured	small
U_L	unstructured	large
U_S	unstructured	small

- *Instantaneous Self Rating (ISA) Scale:* The workload rating the participant indicates during an instance of time throughout the scenarios,
- *Correlation DD-TBX:* Correlation of DD with TBX to see how similar or dissimilar the metrics are in a 4D trajectory management environment and for this new sector,
- *Correlation TBX-ISA:* Correlation to verify how close the TBX metric can predict workload which will give a better indication of the actual inherent complexity,
- *Correlation DD-ISA:* Correlation to verify how close the DD metric can predict workload which will give a better indication of the actual inherent complexity.

F. Scenarios

Participants were asked to reroute air traffic safely in a large hypothetical en-route sector ($\approx 230,000 km^2$) under the four different control conditions shown in Table II. To fully test the complexity metrics capabilities the large sector was designed in order for more entry/exit waypoints (18) to create numerous initial aircraft headings and routes for an extremely structured baseline case and an extremely unstructured baseline case as well as room for more LOS in multiple locations to simulate a future 4D trajectory management environment (i.e., higher aircraft capacity and big sector area responsibility). The rotation of the boxy shaped sector varied ($0°$ or $180°$) between scenarios consisting of the same (baseline) traffic structure to avoid a control bias due to scenario recognition. Additionally, the names of the entry/exit waypoints and aircraft call signs were also varied for each scenario to prevent this bias. An example of what the sector looks like and of structured traffic with small perturbation (S_S) is shown in Fig. 2. Notice the initial heading of two traffic streams (top and bottom right) are pointed directly at the restricted area and the four streams create a clockwise traffic around the restricted area. An example of the unstructured traffic with large perturbation (U_L) is shown in Fig. 3. In this example an aircraft from the top left corner is causing the multiple conflict warnings with aircraft that was previously routed safely.

In each scenario 52 aircraft were presented to the participant and lasted 60 minutes in scenario-time. The simulation ran at four times the normal speed, such that each scenario lasted 15 minutes in real-time. Using data from a prior experiment it was determined that for sufficient workload resolution an ISA scale would pop-up with an audible tone with the interval of 100 seconds scenario-time (25 seconds real-time). Participants were instructed to quickly and accurately indicated their workload on this scale from 0 to 100 that would disappear once they entered their rating each time. Each scenario was designed to have a build up of aircraft with the peak of around 45 aircraft in the sector at the 30 scenario-minute mark, halfway. After the halfway mark the aircraft entering dropped off considerably till the end. Although, because of the large sector area most of the aircraft were still lingering (large aircraft count) in the sector and eventually reaching their exit waypoints. All aircraft entered the controlled sector at FL300 through one of eighteen fixed waypoints on the sector border with their initial (straight) 4D trajectory pointing towards one of the other waypoints. The aircraft could only be controlled laterally (i.e., vertical manipulations of the trajectories was not possible), and only if they were physically inside the sector. Nonetheless, aircraft inbound to the sector where shown in gray when approaching such that the participants had ten minutes (scenario-time) to prepare for future traffic situations. All aircraft had the same simulated performance characteristics of a single generic aircraft type.

Although initial conditions of each scenario were set such that the controller had to resolve a certain amount of perturbations (conflicting pairs of aircraft and avoiding restricted areas) by manipulating the trajectories of individual aircraft, the control actions themselves could create new conflicts and restricted area intrusions further ahead in time. During the development of the scenarios with the Travel Space tool, it resulted in the tool itself being improved to become more intuitive with continuous situational awareness making ISA intervals of 25 seconds real-time achievable.

G. Hypothesis

It was hypothesized that the TBX would predict workload better than DD in a 4D trajectory management environment. Because TBX is trajectory-based it should align well with the workload ratings given by participants. The design of the experiment made it so that there would be a big sector giving controllers the selection of more rerouting solutions using the Travel Space tool thus allowing all participants to safely control the traffic and resolve all perturbations. It was also hypothesized that because DD is state-based it would not correlate well with the workload ratings because of the close spacing and trajectories of aircraft that could appear to be in a collision course from just looking at an instance of time, when in fact their future trajectories would not meet at all. Because of the advantages of 4D trajectory-based tools such as constant situational awareness, controllers were predicted to more likely space aircraft closer together with confidence and rely heavily on the tool. Lastly, it was predicted that the large perturbation scenarios in both the structured and unstructured cases would prove more difficult because of reduced usable airspace for rerouting causing narrowing of the travel space solutions if there are many waypoints to reroute an aircraft.

H. Data Analysis

A Friedman test was run to determine if there were differences in participants ISA scoring (workload) during varying conditions of traffic structure and perturbation levels. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. Statistical significance was accepted at

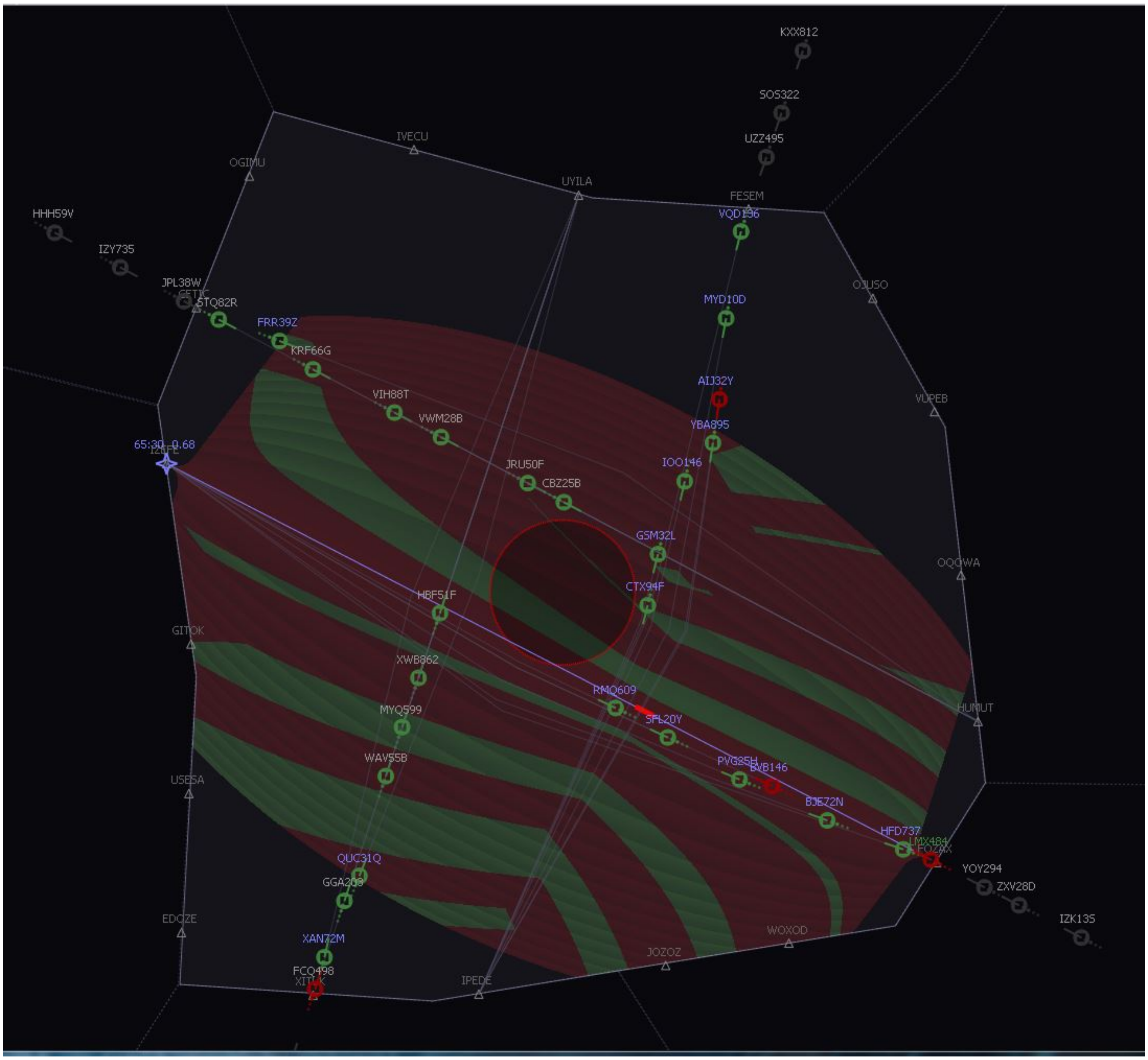


Fig. 2. S_S - Structured traffic with small perturbation.

TABLE III
SIGNIFICANT DIFFERENCES IN PARTICIPANTS ISA SCORING

Comparison between scenario conditions	
$\chi^2(3) = 28.875$	$p < .01$
S_S ($Mdn = -0.57$) to U_L ($Mdn = 0.58$)	$p < .01$
S_S ($Mdn = -0.57$) to U_S ($Mdn = -0.09$)	$p = .010$
S_S ($Mdn = -0.57$) to S_L ($Mdn = 0.10$)	$p = .016$

the $p < .0125$ level. Complexity metrics were plotted against ISA scoring results to observe any trends and correlations. This revealed time shifts in the TBX metric.

VII. RESULTS

Participants ISA scoring was statistically significantly different at the different scenario conditions during the Travel Space experiment runs, $\chi^2(3) = 28.875$, $p < .01$. Post-hoc analysis revealed statistically significant differences in participants ISA scoring as shown in Table III. There was no significant differences in scenario conditions S_L to U_S, S_L to U_L, and U_S to U_L. In the following sections the effect of traffic structure and perturbation on each measure will be discussed.

A. Loss of Separation

Out of the 3328 controlled flights, one loss of separation occurred during the unstructured large perturbation scenario (U_L). This was a result of the participant unnecessarily rerouting aircraft that had no conflicts in order to anticipate future LOS and give aircraft more spacing which distracted the participant from noticing new aircraft entering the sector. This often triggered multiple LOS warnings which resulted in high workload for the participant searching and resolving the pairs,

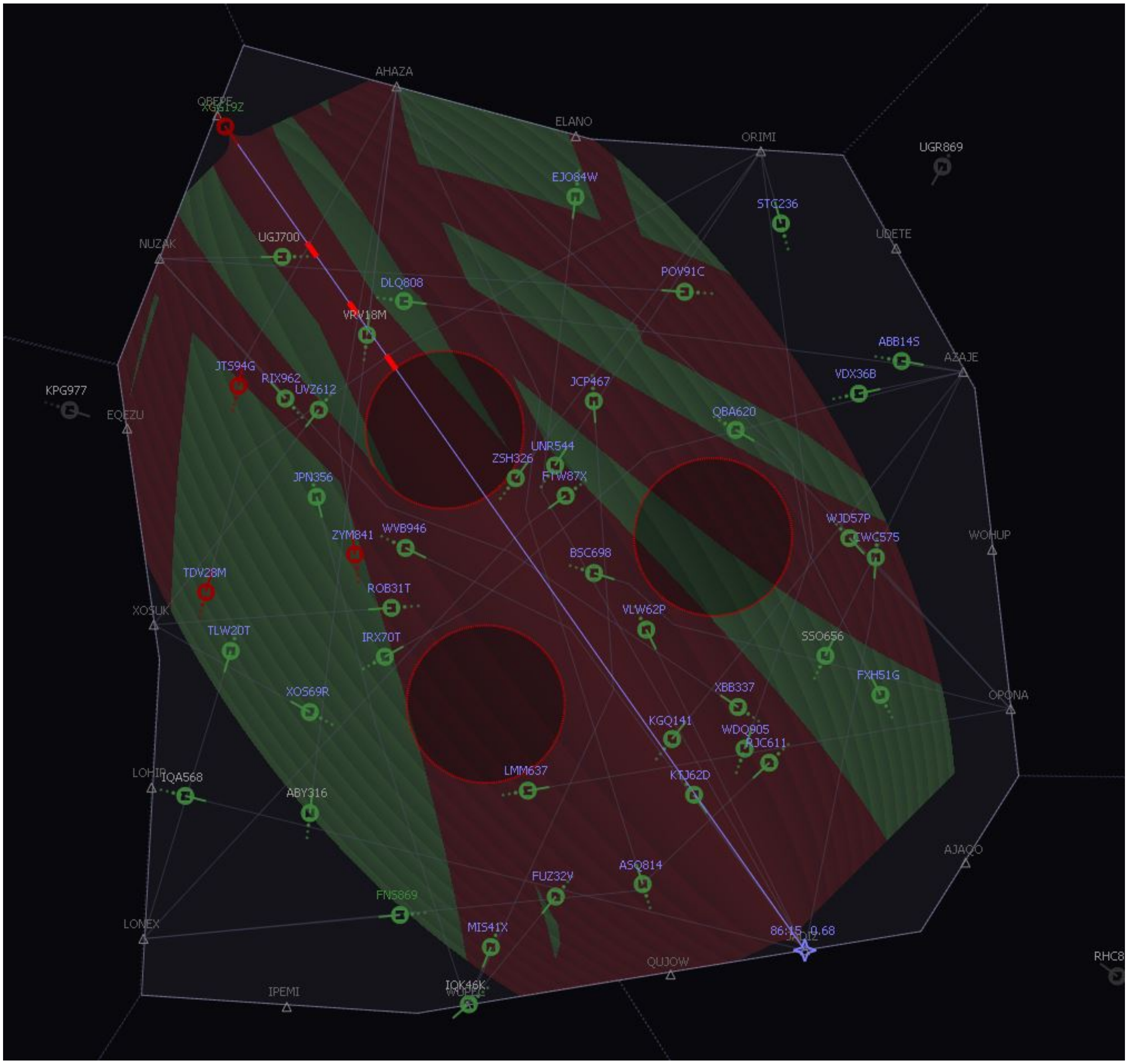
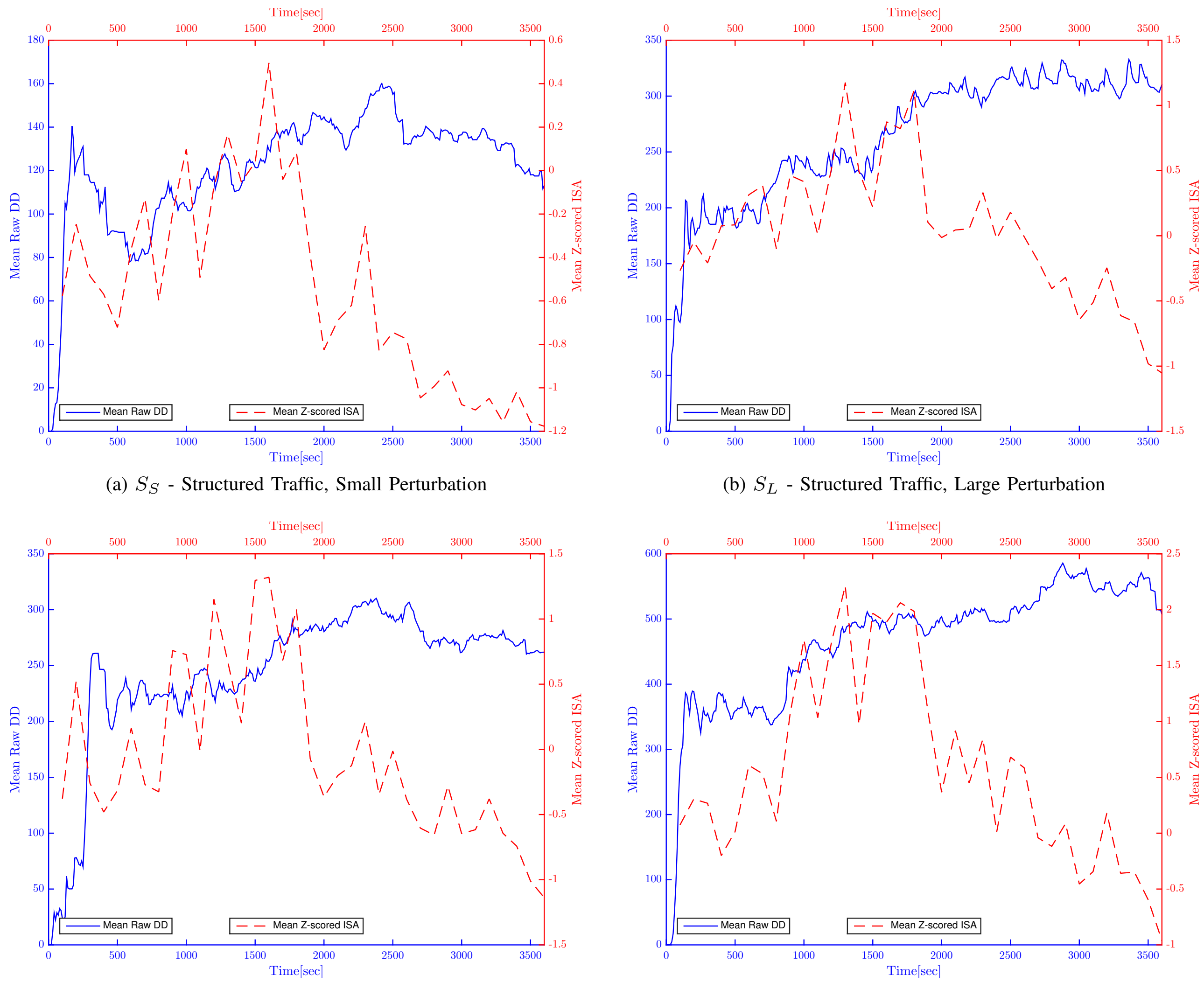


Fig. 3. U_L - Unstructured traffic with large perturbation.

not realizing if one would start by resolving the latest aircraft that entered the sector it would have solved most, if not all of the LOS warnings. The result was the participant did not have enough time to search and resolve a pair of conflicting aircraft which had the latest aircraft entering the sector.

B. Dynamic Density

Fig. 4 shows the plots of the mean values of DD and ISA over all participants for each scenario condition. The values of DD throughout all plots jump up and then steadily increases slowly with only the small perturbation scenarios (Fig. 4a and Fig. 4c) slightly decreasing at around 2500 seconds. With regards to the traffic structures effect on the DD it can be seen that for unstructured air traffic the values are significantly higher throughout the runs than structured air traffic scenarios. For unstructured scenarios the average DD value in Fig. 4c is 244 and in Fig. 4d is 464, compared to structured scenarios average values of 122 in Fig. 4a and of 263 in Fig. 4b. Notice it is about doubled, $244/122 = 2$ and $464/263 \approx 1.8$. With regards to the perturbation effect it can also be clearly seen that just as with traffic structure, it too roughly doubles with higher perturbation. Comparing large and small perturbation in structured conditions the average values are $263/122 \approx 2.2$ and in unstructured conditions $464/244 \approx 1.9$. Overall, the DD metric suggests that unstructured traffic doubles complexity and adding large perturbations doubles the complexity again. This indicates the notion that workload experienced should be constantly very high for the human controller throughout the runs.



(a) S_S - Structured Traffic, Small Perturbation
(b) S_L - Structured Traffic, Large Perturbation
(c) U_S - Unstructured Traffic, Small Perturbation
(d) U_L - Unstructured Traffic, Large Perturbation

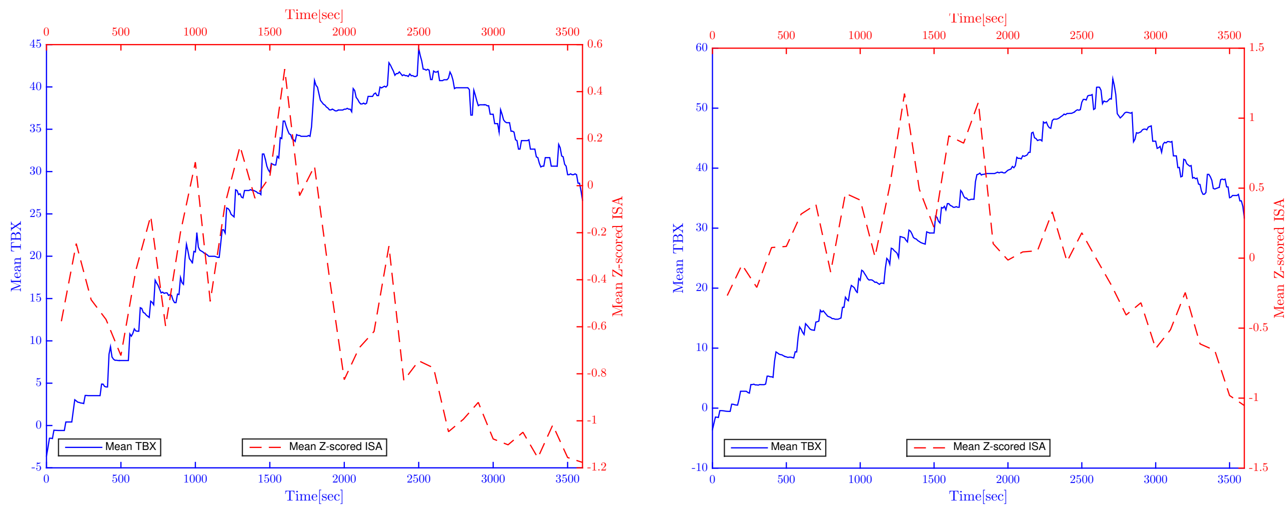
Fig. 4. Mean values of Dynamic Density Metric vs mean Z-scored ISA ratings over all participants.

C. Trajectory-based Complexity

Fig. 5 and Fig. 6 show the plots of the mean values of TBX and ISA over all participants for structured traffic and unstructured traffic respectively. In both figures the values of TBX throughout all plots show a bell shape curve with their peaks roughly around 2500 seconds. With regards to the traffic structures effect on the TBX it can be seen that there is little difference between structured and unstructured scenarios; only the in the large perturbation conditions the structured scenario (Fig. 5b) has a peak value of 54.87 whereas unstructured scenario (Fig. 6b) has a peak value of 47.16 (difference of 7.71). With regards to perturbation effect it can be seen that the large perturbation condition increases the TBX value. In structured scenarios the peak value is 54.87 for large perturbation (Fig. 5b) and a peak value of 44.55 for small perturbation (Fig. 5a), giving a difference of 10.32. For unstructured scenarios the peak value is 47.16 for large perturbation (Fig. 6b) and a peak value of 44.87 for small perturbation (Fig. 6a), giving a difference of 2.29. Note that the structured traffic and large perturbation scenario (S_L) had the biggest peak difference out of all scenarios, although only 7.71 and 10.32 respectively. This indicates that workload experienced will be roughly the same in all scenarios and the different levels of traffic structures or perturbations do not affect the human controllers main task of safely controlling aircraft too much in a TBO environment.

D. Instantaneous Self Rating Scale

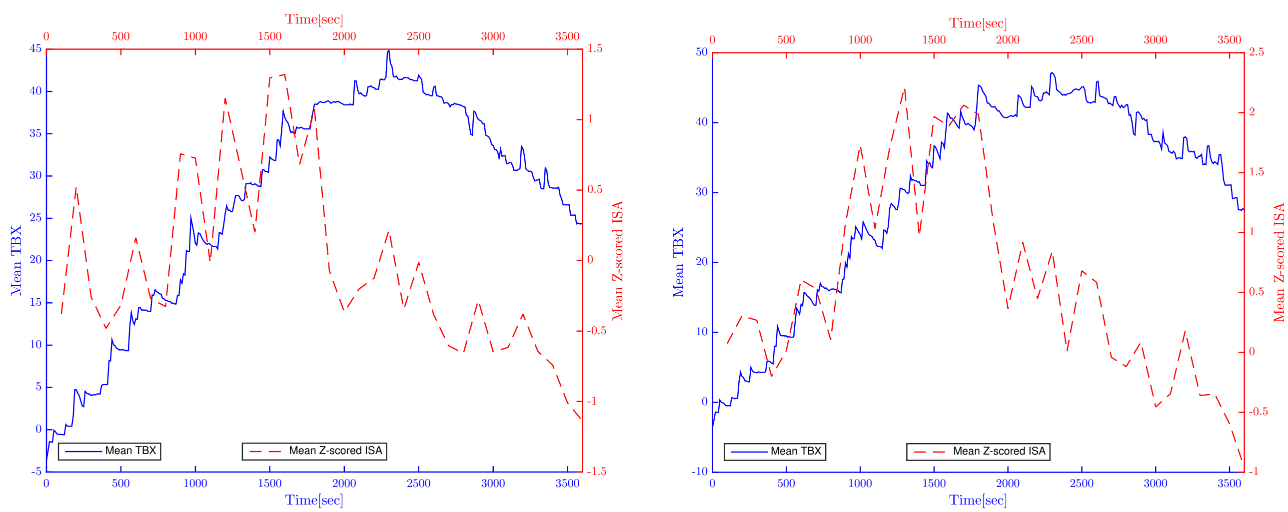
Comparison plots of the mean ISA scores over all participants are shown in Fig. 7 and Fig. 8. The ISA scores are Z-scored per participant. In all scenarios we see the general trend of the workload rating increasing until roughly around the 1500 second mark (almost halfway into experiment run) and rapidly decreases, exhibiting a rough skewed bell curve shape. With regards to the traffic structures effect on the ISA scores it can be seen in Fig. 7 that unstructured air traffic significantly increases the workload on average both in small and large perturbations conditions. The unstructured and structured mean of the difference is around 0.5 for both perturbation conditions. In Fig. 7a the biggest difference is



(a) S_S - Structured Traffic, Small Perturbation

(b) S_L - Structured Traffic, Large Perturbation

Fig. 5. Structured scenarios mean values of Trajectory-based Complexity Metric verse mean Z-scored ISA ratings over all participants.



(a) U_S - Unstructured Traffic, Small Perturbation

(b) U_L - Unstructured Traffic, Large Perturbation

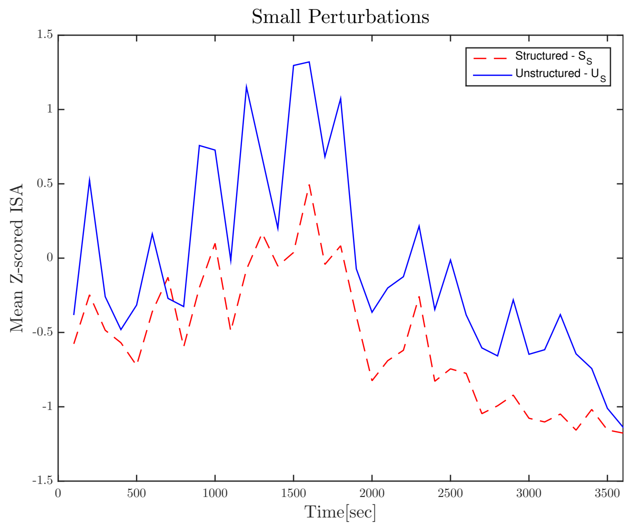
Fig. 6. Unstructured scenarios mean values of Trajectory-based Complexity Metric vs mean Z-scored ISA ratings over all participants.

1.26 with the highest peak of 1.32 in U_S and in Fig. 7b the biggest difference is 1.75 with the highest peak of 2.21 in U_L. With regards to the perturbation effect it can be seen in Fig. 8 that large perturbation significantly increases workload both in structured and unstructured traffic conditions. The large and small perturbations mean of the difference is around 0.6 for both traffic structure conditions. In Fig. 8a the biggest difference is 1.03 with the highest peak of 1.18 in S_L and in Fig. 8b the biggest difference is 1.54 with the highest peak of 2.21 in U_L. This indicates that participants are significantly experiencing higher workloads with unstructured traffic and/or larger perturbations conditions in the TBO environment.

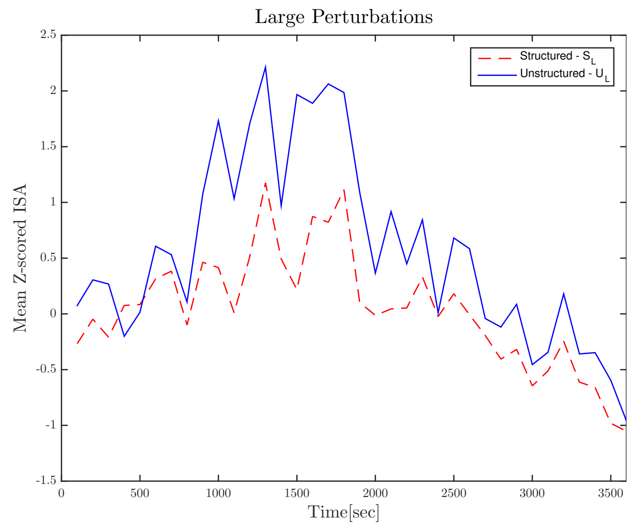
E. Correlation DD-TBX

Due to DD coefficients only being obtained empirically after numerous runs of a particular air sector in a state-based manner, it is only possible to compare the raw score elements of DD individually to the TBX for this experiment study. Because the two metrics work on different principles (state-based vs trajectory-based) it is more interesting to see at which elements they differ the most under the different scenario conditions in a TBO environment (if both metrics correlated perfectly there would be no use for a comparison study). The results of the DD elements varied between all scenarios with the similar level of correlation results to slight variation and only a few with very noticeable difference.

In Fig. 9 the largest difference is shown in the Horizontal Proximity Metric (HOR) elements of DD with TBX. With regards to the effect of traffic structure on the correlations of DD and TBX it is noticed in Fig. 9a that the biggest difference is that structured conditions have very little correlation (≈ 0.08) while the unstructured conditions have very high correlations (≈ 0.93). In Fig. 9b similar results can be seen although the structured conditions have a good amount of correlations (0.55) and unstructured conditions with very high correlations (≈ 0.96). In Fig. 9c notice the same trend of unstructured



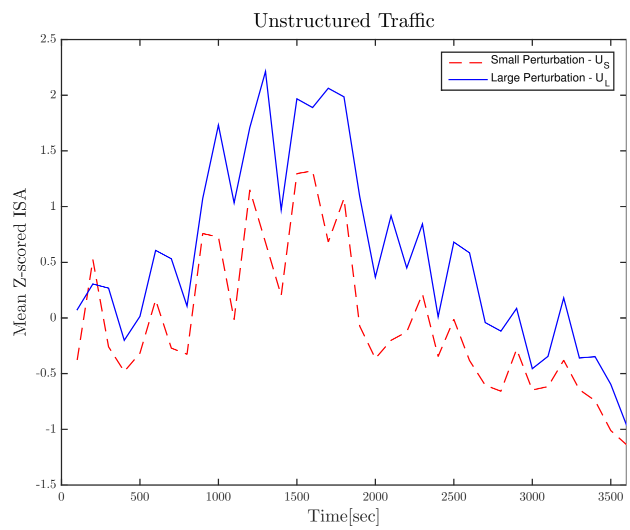
(a) Traffic effect in small perturbation condition



(b) Traffic effect in large perturbation condition

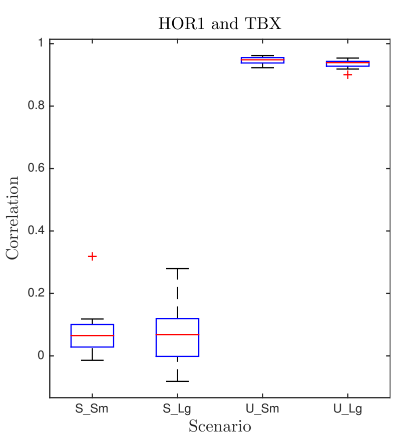
Fig. 7. Mean Z-scored ISA ratings traffic structure comparison.

(a) Perturbation effect in structured traffic condition

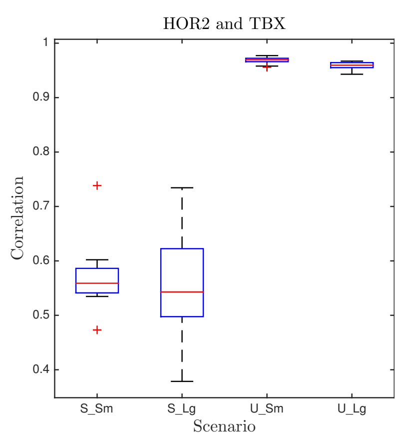


(b) Perturbation effect in unstructured traffic condition

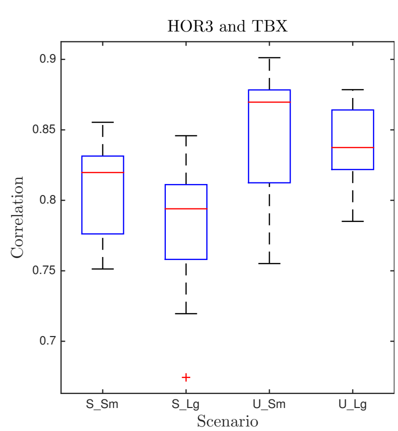
Fig. 8. Mean Z-scored ISA ratings perturbation level comparison



(a) Correlation of DD HOR1 with TBX



(b) Correlation of DD HOR2 with TBX



(c) Correlation of DD HOR3 with TBX

Fig. 9. Correlation of HOR elements of Dynamic Density Metric with the TBX Metric

conditions having higher correlations is shown although not as strong. With regards to the effect of perturbation levels it can be seen in Fig. 9 that all cases exhibit small perturbations having a higher on average correlation than large perturbations, in all traffic structures. However, this appears to be slight but noticeable. These trends are explained when the aircraft enter the sector and depending on how close they are with all other aircraft the entire sectors aircrafts' average distant to one another gets calculated. In structured traffic the average distances will be greater as traffic are coming in streams from four specific entry waypoints that are spaced out far when they enter the sector and exit the sector at four exit waypoints. For unstructured traffic the aircraft come from all 18 waypoints of the sector which will spread out aircraft the moment they enter so the average distances of aircraft to one another will be lower. In other words, the reason for strong correlation in unstructured conditions is when more aircraft enter the sector the average distances between all aircraft gradually reduces giving a higher HOR value, and as TBX is essentially a modified aircraft count it gradually goes up as more aircraft enter the sector.

F. Correlation DD-ISA

From Fig. 4 it shows that there are no similarities of the DD to the workload experienced indicated from ISA scores; the DD scores increase steadily till the full run while the ISA scores exhibit a bell curved shape. This can be confirmed from Fig. 10 where correlation plots were calculated for each element of DD to ISA scores. The DD elements are taken separately as no extensive runs (controlled in a state-based manner) of the experiment sector has been done in order to empirically determine the DD coefficients for each element. All the results have varying or very little correlation trends so not much can be said about the effect of either traffic structure and perturbation levels. For HOR1 and HOR2 there seems to be higher correlations for structured conditions however this is very small and the large perturbations conditions have a range of slightly negative to about 0.25. This was to be expected because DD looks at a current state of traffic, whilst in fact the complete scenario could have already been managed such that it is conflict free. Additionally, these are the raw DD scores that are not weighted so the coefficients could had very different effect or weightings on the full DD score. This was expected as DD is state-based calculating complexity in a TBO environment where workload experienced is much different with a strategic management approach rather than a 'hands-on' approach that the metric was initially designed around.

G. Correlation TBX-ISA

In the original plots seen in Fig. 5 and Fig. 6, which have the TBX plotted with ISA scores, it can be seen to have a clear trend throughout all scenario conditions. However, there appears to be a time shift where the TBX responds later on in time than the participants current workload that they indicated using the ISA scale; TBX is essentially a modified aircraft count. This explains the initial negative, zero, or low correlations seen in Fig. 13a. A time shift was applied to each scenario condition and as seen in Fig. 11 and Fig. 12 all have similar trends.

With regards to the effect traffic structure has on the correlation of TBX and ISA it can be seen by comparing Fig. 11a to Fig. 12a and comparing Fig. 11b to Fig. 12b that the time shift in the unstructured conditions are 100 and 200 seconds lower for small and large perturbations respectively. This suggests that the human controller has more time planning of safely controlling traffic further ahead in time when the air traffic is structured than unstructured. This was observed in participants as they noticed traffic only coming from four entry waypoints and formulated a strategy which they stuck with to navigate through restricted areas throughout the experiment run, only deviating slightly when an aircraft did not fit within their traffic patterns. In Fig. 13b it can also roughly be seen that unstructured conditions have a higher correlation than structured conditions in addition to all scenarios having positive correlations after the time shifts. An observation that explains this is when traffic was predictable (structured), participants workload dropped dramatically where they would score very low ratings because they seemed very confident about future time events and handling new aircraft entering the sector, except for a surprise which would jump their rating high again briefly. With unpredictable (unstructured) traffic participants were on their game to find any impending LOS or restricted area intrusions and appeared to have less time to resolve conflicts as aircraft entering anywhere in the sector were difficult to spot if they were busy resolving another aircraft pair or multiple conflicts. In other words, their higher workload ratings in unstructured traffic would not drop as dramatically as in structured traffic (giving a closer trend with TBX) because they will need to search where the new aircraft entering the sector came from that is causing instant multiple aircraft/restricted area conflicts; solving the latest aircraft will solve multiple conflicts (killing two birds with one stone).

With regards to the effect of perturbation levels on the correlations of TBX and ISA it can be seen that large perturbation has higher correlations although a wider spread than smaller perturbations as seen in Fig. 13b. In Fig. 11 and Fig. 12 it can be seen that the large perturbations have 200 and 100 seconds longer time shifts for structured and unstructured traffic respectively. In experiment observations the large perturbation condition for both traffic conditions meant more forward planning for the controller in order to compensate for the reduced airspace to reroute traffic in; planning wisely with fewer travel space options. The higher correlation of large perturbation with workload than with small perturbation is due to planning ahead and being alert to conflicts as the usable airspace is reduced for safe air traffic control, as such their workload remained relatively higher throughout because of their attentiveness; resembling TBX more so as its a modified aircraft count that is gradual.

H. Debrief

After the experiment participants were invited to comment on the effectiveness of the Travel Space tool and their strategies using the tool. All participants reported that the tool

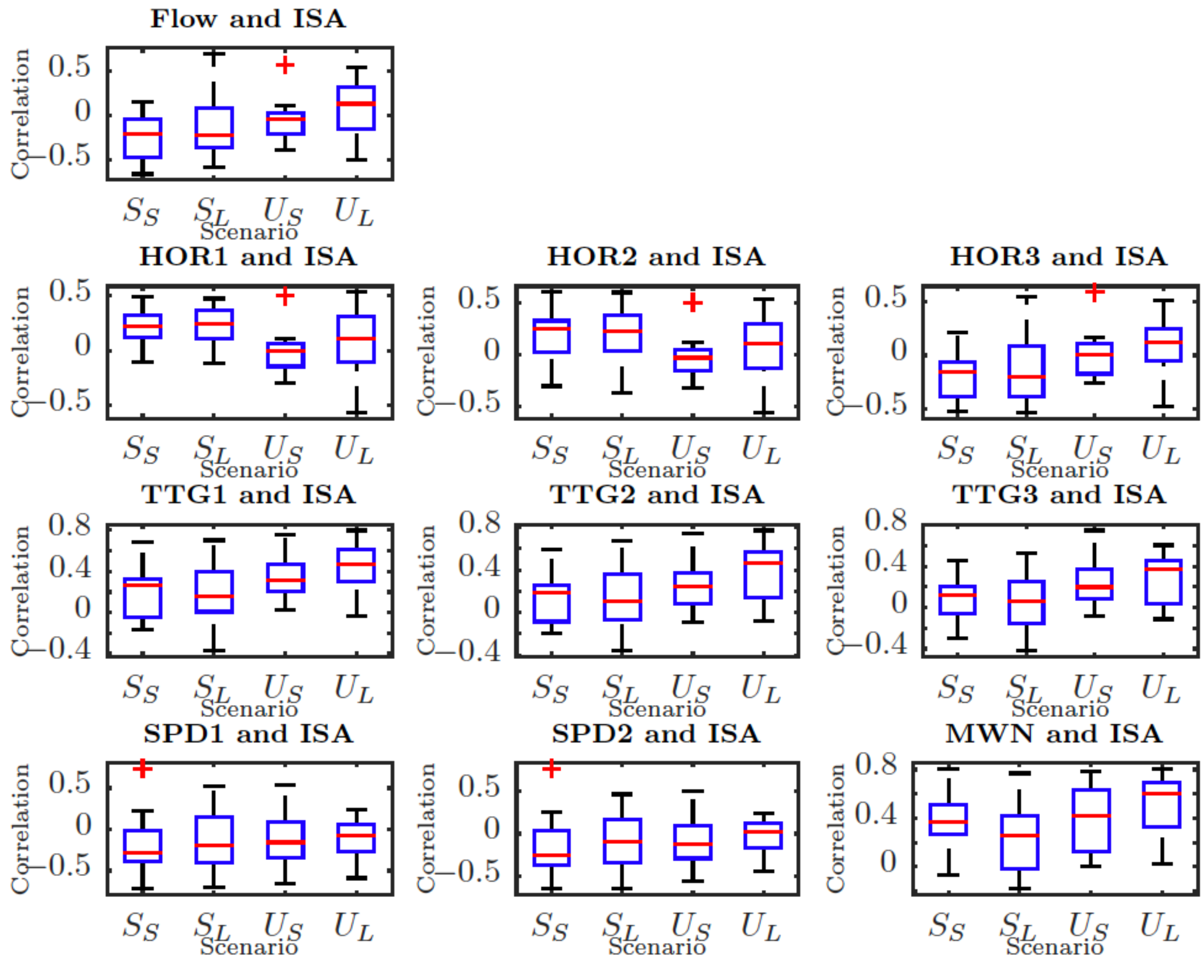
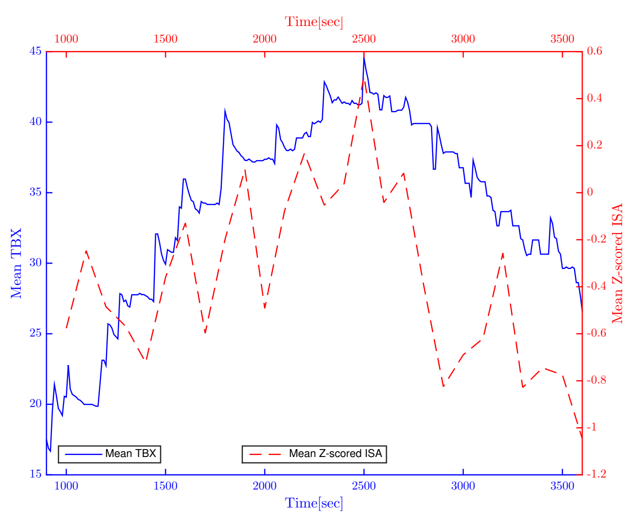
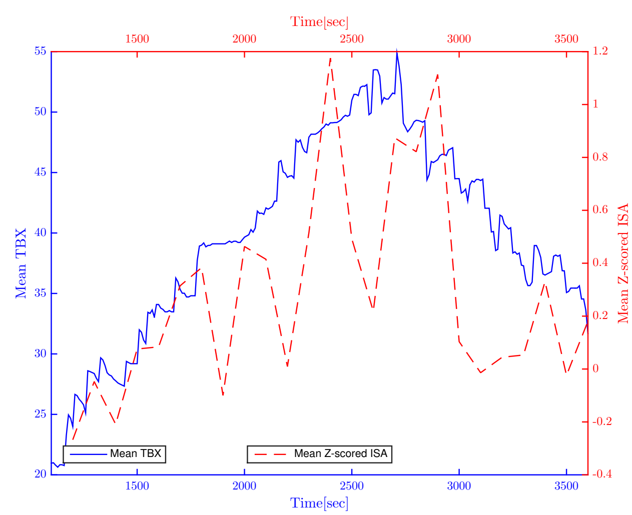


Fig. 10. Correlation of Dynamic Density Metric with ISA scores



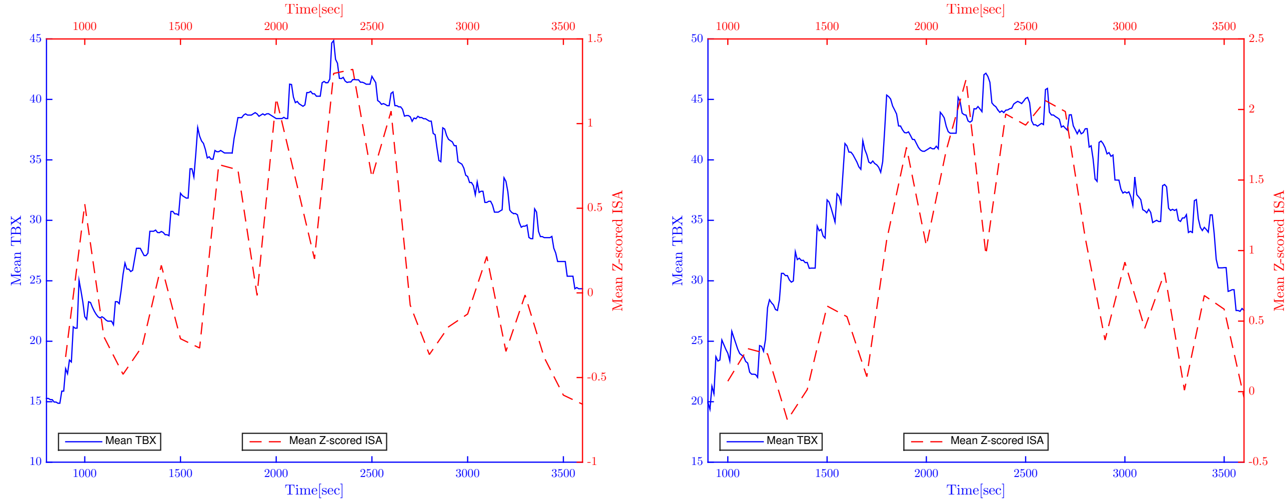
(a) S_S - Time shift of 900 seconds



(b) S_L - Time shift of 1100 seconds

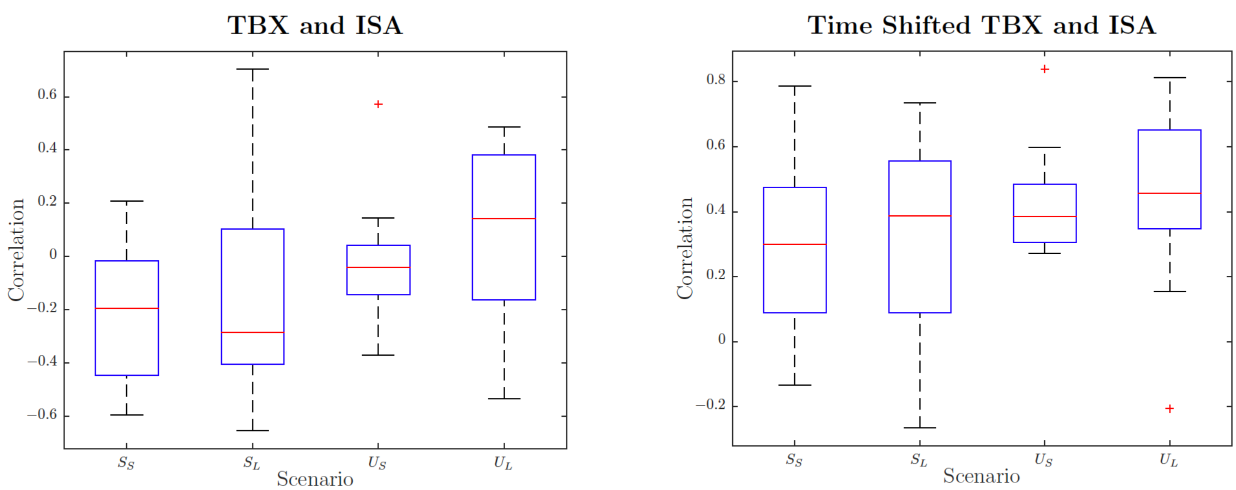
Fig. 11. Structured scenarios time shifted mean ISA scores verse mean values of Trajectory-based Complexity Metric over all participants.

was intuitive and provided clear representations of solutions to safely rerouting a selected aircraft. Most participants mentioned that they relied heavily on the tool to visualize solution options for rerouting traffic safely, disregarding additional information aids on each aircraft or the LOS warning list. Some commented the addition information (e.g., exit waypoint, speed) on each aircraft was useless because all aircrafts' entire trajectories were visually shown throughout the runs



(a) U_S - Time shift of 800 seconds (b) U_L - Time shift of 900

Fig. 12. Unstructured scenarios time shifted mean ISA scores verse mean values of Trajectory-based Complexity Metric over all participants.



(a) Correlation of TBX with ISA (b) Correlation of Timeshifted TBX with ISA

Fig. 13. Correlations of Trajectory-based Complexity Metric with ISA scores

by gray lines (continuous situational awareness). Only one participant reported to briefly use the LOS warning list for their first scenario run. Many participants mentioned that only later on they figured out that once the aircraft in the sector were safely controlled with no conflicts, new or multiple conflicts were caused by new aircraft entering the sector to which it would be best to reroute the new aircraft first before trying to reroute other aircraft, saving valuable time. Some mentioned it was difficult seeing how the green zones (e.g., safe field of travel) were obtained because depending on where one places aircraft it would affect each other aircraft's Travel Space options. In some instants participants did not see any green zones for the selected aircraft and were franticly moving other aircraft to find green zones for all aircraft. A suggestion was for the tool to have the option to show up all aircrafts' Travel Space who were in conflict with the selected aircraft as well in order to better visualize the best solution spaces for aircraft at the same time. A common complaint was the frequency of the ISA window popping up for them to score (every 25 seconds in real-time) which was a nuisance and at some instances, while rerouting multiple traffic, adding to the workload experienced. Participants were not told that the aim of the experiment was measuring their workload while performing under various conditions with the tool. However, despite their perceived nuisance of the ISA timing interval, it was necessary for the resolution desired to compare the complexity metrics to while not giving too much of a burden to participants.

VIII. Discussion

The use of complexity metrics in 4D trajectory management environments have shown to be promising in predicting human controller workload during various traffic orderliness and perturbation levels. The implications of being able to predict workload of a new airspace sector and/or a new 4D trajectory management tool is paramount for future designs of safer travel with higher capacity.

The results of the human-in-the-loop experiment showed that the TBX metric is promising to predict complexity within a sector. The time shift of the TBX to actual experienced workload shows that in the designing of 4D trajectory management tools this could prove to be useful to see how far in advance the workload has been shifted, providing a bigger safety window for controllers to safely manage traffic. Although participants had no operational controller experience it shows how intuitive 4D trajectory management tools can or should be as the shift of ATCOs are towards strategic management approaches in the future. It was observed that students often opted for tighter routing of traffic around restricted area edges to free airspace for future maneuverability of aircraft instead of routing traffic freely around restricted areas; this pushes the reliability, accuracy, and precision of the tool to greater importance for aircraft safety. If a controller were to doubt the decision supporting tool it could revert back to 'hands-on' control which could prove detrimental and likely impossible to handle in a large 4D trajectory management environment.

IX. Conclusion

The goal of the study has been to empirically investigate if the DD metric and/or the TBX metric can predict human controller workload in a 4D trajectory management environment. In an experiment the initial traffic orderliness and the level of perturbation have been manipulated. The human-generated control actions determined the calculations of the two complexity metrics being studied as well as affecting their level of experienced workload they reported. Results show that the TBX metric roughly represents the controller workload although time shifted. This proves promising that the TBX metric could predicted controller workload in 4D trajectory management, aiding in designing safer future airspace sectors.

References

[1] "SESAR Concept of Operations," Brussels, Belgium, Report, 2007.

[2] "Concept of Operations for the Next Generation Air Transportation System, Version 3.2," Joint Planning and Development Office, Report, 2011.

[3] B. Hilburn, "Cognitive Complexity in Air Traffic Control: A Literature Review," Center for Human Performance Research, Tech. Rep., Feb. 2004.

[4] L. Bainbridge, "Ironies of Automation," *Automatica*, vol. 19, no. 6, pp. 775–779, Nov. 1983.

[5] G. Baxter, J. Rooksby, Y. Wang, and A. Khajeh-Hosseini, "The ironies of automation... still going strong at 30?" in *Proceedings of the 30th European Conference on Cognitive Ergonomics - ECCE '12*. New York, New York, USA: ACM Press, 2012, pp. 65–71.

[6] M. R. Endsley and D. B. Kaber, "Level of automation effects on performance, situation awareness and workload in a dynamic control task," *Ergonomics*, vol. 42, no. 3, pp. 462–492, 1999.

[7] R. Parasuraman, T. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 30, no. 3, pp. 286–297, May 2000.

[8] K. Christoffersen and D. D. Woods, "How to Make Automated Systems Team Players," in *Advances in Human Performance and Cognitive Engineering Research*. Elsevier Science Ltd., 2002, vol. 2, ch. 1, pp. 1–12.

[9] J. M. Flach, F. Tanabe, K. Monta, K. J. Vicente, and J. Rasmussen, "An Ecological Approach to Interface Design," *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, vol. 42, no. 3, pp. 295–299, Oct. 1998.

[10] D. A. Norman, "The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'." *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, vol. 327, no. 1241, pp. 585–93, Apr. 1990.

[11] K. J. Vicente and J. Rasmussen, "The Ecology of Human-Machine Systems II: Mediating 'Direct Perception' in Complex Work Domains," *Ecological Psychology*, vol. 2, no. 3, pp. 207–249, Sep. 1990.

[12] T. Prevot and P. U. Lee, "Trajectory-Based Complexity (TBX): A Modified Aircraft Count to Predict Sector Complexity During Trajectory-Based Operations," in *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*. Seattle, Washington, USA: IEEE, Oct. 2011, pp. 3A3–1–3A3–15.

[13] P. Kopardekar and S. Magyarits, "Dynamic Density: Measuring and Predicting Sector Complexity [ATC]," in *Proceedings. The 21st Digital Avionics Systems Conference*, vol. 1. IEEE, 2002.

[14] M. M. van Paassen, C. Borst, R. Klomp, M. Mulder, P. van Leeuwen, and M. Mooij, "Designing for shared cognition in air traffic management," *Journal of Aerospace Operations*, vol. 2, no. 1-2, pp. 39–51, 2013.

[15] G. B. Chatterji and B. Sridhar, "Measures For Air Traffic Controller Workload Prediction," *Proceedings of the 1st AIAA, Aircraft, Technology Integration, and Operations Forum*, Oct. 2001.

[16] S. M. B. Abdul Rahman, M. Mulder, and M. M. van Paassen, "The Effects of Air Traffic Control Sector Design on the Solution Space Diagram," in *Proceedings of ECCE 2010 Conference*, W.-P. Brinkman and M. Neerincx, Eds., European Association of Cognitive Ergonomics. Mediamatica, Delft University of Technology, The Netherlands, Aug. 2010, pp. 261–268.

Jason Toy received the B.S. degree (2007) in aerospace engineering from the University of California, San Diego in La Jolla, The United States of America. He is currently working towards the M.Sc. degree in aerospace engineering at Delft University of Technology, Delft, The Netherlands at the section Control and Simulation. His research focuses on investigating measuring complexity in ecological 4D trajectory-base controller decision support tools. His interests include human-machine systems, ecological interface designs, and research simulators.

Clark Borst received the M.Sc. (2004, with honours) and Ph.D. degrees (2009) in aerospace engineering from the Delft University of Technology, Delft, The Netherlands, where he is currently working as an Assistant Professor in aerospace human-machine systems. His work and research interests include developing human-centered aviation automation through the application and empirical evaluation of Cognitive Systems Engineering and Ecological Interface Design principles.

Rolf Klomp received the M.Sc. degree (2010) in aerospace engineering from the Delft University of Technology, Delft, The Netherlands, where he is currently working towards the Ph.D. degree at the section Control and Simulation. His research focuses on the development of ecological gls4D trajectory-based controller decision support tools. His research interests include human-automation interaction, air traffic simulation and (ecological) interface design.

Max Mulder received the M.Sc.degree and Ph.D. degree (cum laude) in aerospace engineering from the Delft University of Technology, Delft, The Netherlands, in 1992 and 1999, respectively, for his work on the cybernetics of tunnel-in-the-sky displays. He is currently Full Professor and Head of the section Control and Simulation, Faculty of Aerospace Engineering, Delft University of Technology. His research interests include cybernetics and its use in modeling human perception and performance, and cognitive systems engineering and its application in the design of ecological human-machine systems.

René van Paassen is an Associated Professor in Aerospace Engineering at the Delft University of Technology, working on human machine interaction and aircraft simulation. He applies Cognitive Systems Engineering analysis (Abstraction Hierarchy, Multi-Level Flow Modeling) and Ecological Interface Design to the work domains of vehicle control and Air Traffic Management. René is also Associate Editor for the IEEE Transactions on Human-Machine Systems.

Appendix A

Research Motivation

A-1 Background

Air Traffic Management (ATM) encompasses all the aviation systems that aid aircraft to depart from an aerodrome, navigate safely through airspaces, and land at a destination aerodrome. Overtime ATM has been evolving with technology and increasing air traffic demands. It is expected that air traffic will more than double over the next couple of decades. (Price, 2011) With the increase of traffic volume it will be inevitable to move towards higher levels of automated support tools.

The systems in aircraft being built today are highly advanced with glass cockpits and modernized sensors. Whereas the ATM system has not seen such a jump in implemented advancements due to different technologies and approaches being developed of how to optimize the system. The issue in ATM is that it has to be able to cope with the entire fleet of aircraft which are currently operational. Aircraft can have a lifespan of approximately 30 years. Therefore, the technology and technological possibilities/implementations lag behind on flight-deck design. Moreover the implementations of upgrading to a newer ATM system takes time as the whole system cannot be switched overnight with zero incidents and with no unknown consequences on top of a constant evolution of the ATM system. It will have to be implemented in stages in order not to bring air traffic to a grinding halt when problems do occur. Despite modern technology advancements it is not possible to fully and accurately model the full complexity of airspace operations. Weather cannot accurately be predicted 100% and unforeseen situations will occur that the automated system have not been programmed to handle. For example, volcano eruptions that produced a massive amount of ash disrupting air traffic, migrating birds, unknown unmanned aerial vehicles, state emergencies to clear airspace, emergency to prioritize select aircraft, etc. Therefore, fully automated Air Traffic Control (ATC) is not possible and thus the human operator is responsible to work together with the automation and monitoring.

The current evolution of the ATM system is foreseen to bring a paradigm shift to the work domain of the Air Traffic Controller (ATCO). The major programs developing future ATM systems are Single European Sky ATM Research (SESAR) in Europe and Next Generation Air Transportation System (NextGen) in the United States of America.

SESAR is a European project that will overhaul the whole European airspace into one so that it is no longer constrained by national boarders and an ATM system to meet future capacity and air safety needs. The project aims to develop a modernized ATM system for Europe that will ensure safety and fluidity of air transport over the next thirty years, making flying more environmentally friendly and reduce the costs of ATM. (Single European Sky ATM Research, 2009)

NextGen is an American project that will transform America's ATC from World War II era technology ground-based system to a satellite-based system to transit the skies. Satellite navigation precision will enable shortened and more direct routes, save travel-time and fuel, reduce traffic delays, increase capacity, and allow controllers to monitor and manage aircraft with greater safety margins. (Federal Aviation Administration, 2013)

In order to cope with the increasing traffic volume it is expected that the evolution will result in a situation where high-precision four-dimensional (4D) trajectories (space and time) for aircraft as a means for strategic management rather than the current -hands on- method of control. In both projects a central role is foreseen for the human operator, aided by higher levels of automation and advanced decision-making tools. As these new ATM systems developed by NextGen and SESAR come on-line gradually in the near future, it can be expected that new problems will emerge as the need for higher level automation and advance decision-making tools will become more prevalent with advancing technology and increasing air traffic demand. It would not be possible for the operator to handle higher traffic capacity without automated support tools as the workload increases drastically with current methods and tools. With more and more automation it is critical to give situational awareness and understanding of what the automated system is doing for the operator to have a mental model of the working environment. Although NextGen and SESAR are developing advanced automation tools it has been shown in the project, Programme for Harmonised ATM Research in EUROCONTROL (PHARE), the lack of trust and reliance the controllers have of automation tools. The irony is that the more we depend on technology and the more advanced it is, the contribution of the human operator being highly-skilled, well-trained, well-practiced is ever more crucial to ensure that the systems are resilient; acting as the last line of defense to save the day against the failures that will inevitably occur. (Bainbridge, 1983; Baxter, Rooksby, Wang, & Khajeh-Hosseini, 2012)

With the current economic troubles in the world and future ones, nations have been slashing budgets across the board. Although ATM systems are vital for civilian and military aviation as well as the economy (i.e., tourism, product trading), politics can play a dirty hand not understanding disastrous consequences if ATM were to suddenly have a cut to any part of it in order to visually "save" money immediately. A recent example is of the 2013 sequestration in the United States of America where mandatory spending cuts forced the Federal Aviation Administration (FAA) to make ATC furloughs and shutdown control towers at small airports that caused air travel delays. By reducing the number of controllers this immediately increased the workload and the potential for fatal errors to occur. Fortunately Congress quickly met within days to fix the FAA budget with a rare bipartisan measure to restore safety, reliability, and confidence of air travel. This highlights the importance and need of making sure the future ATM systems will be able to aid the ATCO in solving the increased air traffic loads with reduced personnel while assuring a high level of air safety and efficiency. Governments around the world want up to seventy five percent less ATCOs which will force the need for 4D trajectory management and ever more highly automated systems. The future scenario of 4D trajectory management will need to address the question of, *"How can we uphold effective human-automation coordination?"*

A-2 Automation and Humans Roles

The greatest limitation of ATM capacity is human controller workload and one of the key factors affecting workload is air traffic complexity.(Hilburn, 2004) It is often believed that because of the limitations of humans, it is better to automate almost everything and in some instances take out the human completely. However, in practice is impossible to eliminated the human entirely for safety reasons. As is inevitable that a failure/situation will occur when the automation does not know how to resolve and relies on a human to solve. Commonly, reactions to evidence of problems in human-automation cooperation have either been one of two directions.(Norman, 1990) Those arguing that these failures are due to inherent human limitations and with more automation we can eliminate the "human error problem." While others arguing that our reach has exceeded our grasp, over-automation is the problem and the solution is to revert to lesser degrees of automated control. Society seems to be locked into thinking that technology and people are independent components; either this electronic box failed or that human box failed.(Christoffersen & Woods, 2002) This is a profound misunderstanding of the factors that influence human performance and that with careful analysis of incidents and disasters many accidents end up representing a breakdown in coordination between technology and people. This leads to the failures being the result of poorly designed automated systems that do not coordinate well with the human operator. The human operator can have a very low situational awareness because the automation is not representing information in a way to see the bigger picture to detect potential problems; despite having lots of information displayed in numbers and dials. When a problem does occur there is high workload experienced because the operator is perplexed not understanding the reasoning behind it and could counter act or override the automation if the trust in the technology is not clearly perceptible.

It is evident that there needs to be an effective human-automation coordination in ATM especially with new types of tools that are being developed which lean heavily on automation. The irony is that the more we depend on technology and the more advanced it is, the contribution of the human operator being highly-skilled, well-trained, well-practiced is ever more crucial to ensure that the systems are resilient; acting as the last line of defense to save the day against the failures that will inevitably occur. (Bainbridge, 1983; Baxter et al., 2012) In human-machine systems an approach called Ecological Interface Design (EID) is a way to tackle this dilemma. This is achieved by looking at the whole working environment and designing systems that make the human operator more situationally aware resulting in having more time to respond to problems and dealing with abnormalities when they do occur. When developing human interfacing tools under EID principles it is important to focus on the analysis of the work domain or environment, rather than on the end user or a specific task. The goal of EID is to make constraints and complex relationships in the work environment perceptually evident to the user; to "make visible the invisible."(Vicente & Rasmussen, 1990;

Flach, Tanabe, Monta, Vicente, & Rasmussen, 1998) This allows the user to devote more time for higher cognitive processes which in the case of ATM would be problem solving and decision making to make better judgments rather than last minute calls.

In order to evaluate a human interfacing tool for its effectiveness there needs to be metrics to analyze it and see if there are correlations to workload. In ATC management tools, the components that need to be measured are the airspace complexity and workload ratings.

A-3 Complexity in ATM

It is believed that by improving measures of ATC complexity it can benefit evaluation of ATM productivity, benchmarking cost effectiveness, assessment of the impact of new tools and procedures, and airspace redesign. Numerous individual components greatly impact the elaborate connection among complexity and workload. Although there are difficulties for fully capturing the notion of cognitive complexity mathematically it does not mean that the human factor must remain unknown.(Hilburn, 2004)

Complexity is defined by the Oxford dictionary as *the state or quality of being intricate or complicated.* (Press, 2013) In ATM complexity is made up of many factors that contribute to the level of complexity for an ATCO. The airspace sector, aircraft count, separation standards, traffic flows, traffic density, aircraft performance characteristics, weather, airspace restrictions, and abnormal events. Other factors come from the tools, displays, arrangements, etc., the ATCO uses.

In ATM there are two types of complexity; inherent and apparent. *Inherent* complexity are the intricate qualities that arise from the airspace properties such as weather, terrain, airspace restrictions, traffic density, etc. *Apparent* complexity are the intricate qualities that come from the interface to the controller. How the information is being displayed to the controller and the tools he/she has and uses to interact within the working environment. Examples would be types of displays (e.g., mono-color, multi-color, touch screens), physical arrangements of displays or consoles, control room layout, software used to display information, etc. The focus of this research will be on the inherent complexity as there is no clear winner of the future ATM system interface to divulge and investigate its apparent complexity.

There are many metrics that have been developed to try and capture airspace complexity fully and needs to be investigated for their relevance to this research experiment. Metrics in the past have focused on state-based complexity metrics. However, the Delft University of Technology (TU Delft) Travel Space tool is trajectory-based and ideally it would be better to analyze it with a trajectory-based complexity metric.

A-4 Research Objective

The objective of this research is to find the equations of airspace complexity that will predict air traffic controller workload in an airspace sector by making an analysis and comparing the elements of Dynamic Density and Trajectory-Based Complexity metrics when using the TU Delft travel space tool in scenario runs with varying scales of traffic orderliness and perturbation levels.

New ATM systems will need to face becoming flexible and dynamic, capable enough to handle unexpected situations. Especially if there is no way that the system can account for these anomalies and will rely on the human controller's judgment to help resolve it. Such anomalies or abnormal events could include weather cells/patterns, hazardous areas (i.e., nuclear disasters, volcanic ash), restricted airspace, aircraft in distress, priority direct routing of special aircraft (i.e., heads of state), beacons not working in sectors, etc. The interface of these systems with air traffic controllers needs to be able to give situational awareness with appropriate levels of workload to the controller in order to solve situations with trust and reliability in the system; giving the controller appropriate information at the right time.

What factors drive controller workload in a 4D trajectory airspace environment?

Currently 4D trajectory management tools are being developed as a part of these new ATC systems which is a new role for the controller as lower level tasks become more automated. The problem facing us today is that no one knows what drives controller workload in a 4D trajectory airspace environment; the particular attributes that contribute to workload in this new way of controlling. However, it is known that a key player affecting workload is air traffic complexity. It is challenging to develop these new trajectory-based systems as it has never been fully tested in the real-world before. It would help immensely saving time, money, and possibly lives if there was a way to analyze and predict the structure of prototype systems and sector areas before even running extensive simulations with ATCOs or real field tests in order to help design an airspace sector.

Is there a way to predict expected workload experienced in an airspace sector with complexity metrics before even running experiments in order to help design sectors?

Looking beyond SESAR and NextGen systems that are currently being developed for the near future it is important to solve problems that will have a high probability of occurring without being bounded by current technologies that are available today, but with technologies predicted to have available in the future. The scope of this research is looking into the future were the technological capabilities, such as hardware infrastructure and detailed data collection from each aircraft, are readily available as to not be bounded by the current or near future hardware limitations or limiting airspace regulations. Rather the focus will be on the interaction between the system automations and the human capabilities (situational awareness, workload, etc.) to being able to optimally interface with the tool even when the tasks become more complex. For simplifications, baby steps, the altitude of aircraft will all be at the same level.

Much effort has been spent on automation, algorithms, optimizers, sensors, etc. However, the role of the human in this system is not yet well defined. Automation tends to work more in a centralized manner despite humans currently controlling air traffic in a decentralized manner by basic means (e.g., voice radio). Furthermore, in EID design, decentralized representations are promoted for giving the controller more situational awareness. The expected direction of future ATM systems are toward high-precision 4D trajectories making the human operator aided more by high levels of automation and advanced decision-making tools. Situational awareness will become increasingly more important to factor in when designing these heavily automated tools. 4D trajectory management is a new role for the human controller and there is a need to design appropriate tools for this future task; automated systems being team players with the human controllers. TU Delft has developed such a tool according to EID principles called the Travel Space tool which will be used in this research as the 4D trajectory testbed when comparing the performance of complexity metrics.

A-5 Research Approach

A 4D trajectory management prototype tool called TU Delft Travel Space was used as a platform for research experiments investigating airspace complexity as future ATM systems will become immersed in 4D trajectory operation environments. The Travel Space tool aided in teasing out the intricate factors that influence human controller workload of a new designed large airspace sector.

Two complexity metrics were analyzed; Dynamic Density (DD) that is state-based and Trajectory-Based Complexity (TBX) that is trajectory-based. Details about these two complexity metrics are discussed in the Appendices C and D on Pages 33 and 43. It was contemplated that for 4D trajectory operations DD might be better in some situations (i.e., structured traffic) where it can act as a state-based/tactical case and in other situations (i.e., weather, emergency, etc.) when flights trajectories are rerouted TBX proves to be a better representation of workload. However, this proved to be difficult to prove because DD needs to develop coefficients in any new sector by means of numerous empirical runs and those coefficients are traditionally obtained when the sector is controlled in a 'hands-on' approach.

The research questions were investigated by conducting a human-in-the-loop experiment using the Travel Space tool in a newly created large airspace sector with increasing inherent complexity levels of perturbation (restricted airspaces) and traffic orderliness (structured and unstructured). The tool in its current form had a two-dimensional (2D) display representation (not showing altitude) with look ahead time, therefore the scope of this experiment was be limited to horizontal trajectories (no vertical manipulations). In other words, altitude changes or vertical crossing angles of aircraft was not considered. During the experiment workload rating measurements were taken in 100 second intervals in scenario-time in which the run was at four times speed, so 25 second intervals in real-time. In designing the airspace sector, traffic orderliness, and perturbation levels it was envisioned to create extreme cases to clearly see which complexity metrics elements correlate to workload, giving each metric a through test.

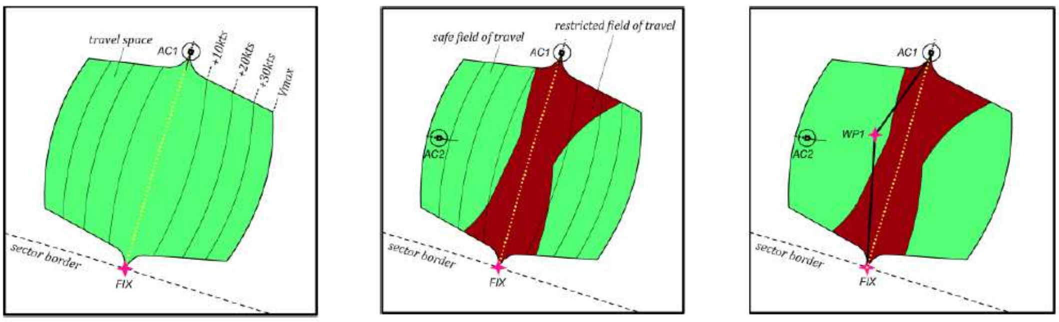
Sixteen participants were tested and all data was collected and analyzed. The two metrics (DD and TBX) were compared against each other using workload measurements obtained from Instantaneous Self Rating (ISA) scores. It was determined that the TBX metric had a time shift in each scenario but otherwise correlated fairly well.

Appendix B

TU Delft Travel Space tool

The Travel Space tool is a 4D trajectory management tool that visually shows the entire planned trajectory path of an aircraft within the sector as well as visually displaying all the possible airspaces that an aircraft can safely maneuver to in the future, based on the aircraft's performance profile. The Travel Space tool was developed at the TU Delft Aerospace Engineering Control and Simulation faculty located in Delft, the Netherlands.

The aim of the Travel Space tool is to assist the ATCO by showing the shared representation of the solution space of a given aircraft to resolve conflicts with a certain look ahead time. The solution space has shapes of ellipses that grow outward with increasing velocity as seen in Fig. B-1(a). The size of the solution space is the limit of what the aircraft is capable of traveling to with its performance. The areas of the solution space are colored green for solutions to conflicts or red for conflict zones which must be avoided as seen in Fig. B-1(b). It is still up to the human controller to select anywhere in the solution space and confirm a trajectory change. In Fig. B-1(c) a placement of a new waypoint is tentatively selected to ensure separation. Once confirmed this creates a new waypoint that will create two new solution space segments. If the original solution space was instead entirely red a new waypoint



(a) Travel space representation for a single aircraft

(b) Traffic causes restricted field of travel (loss of separation)

(c) Placement of intermediate waypoint to ensure separation

Figure B-1: 4D Trajectory selection in Travel Space

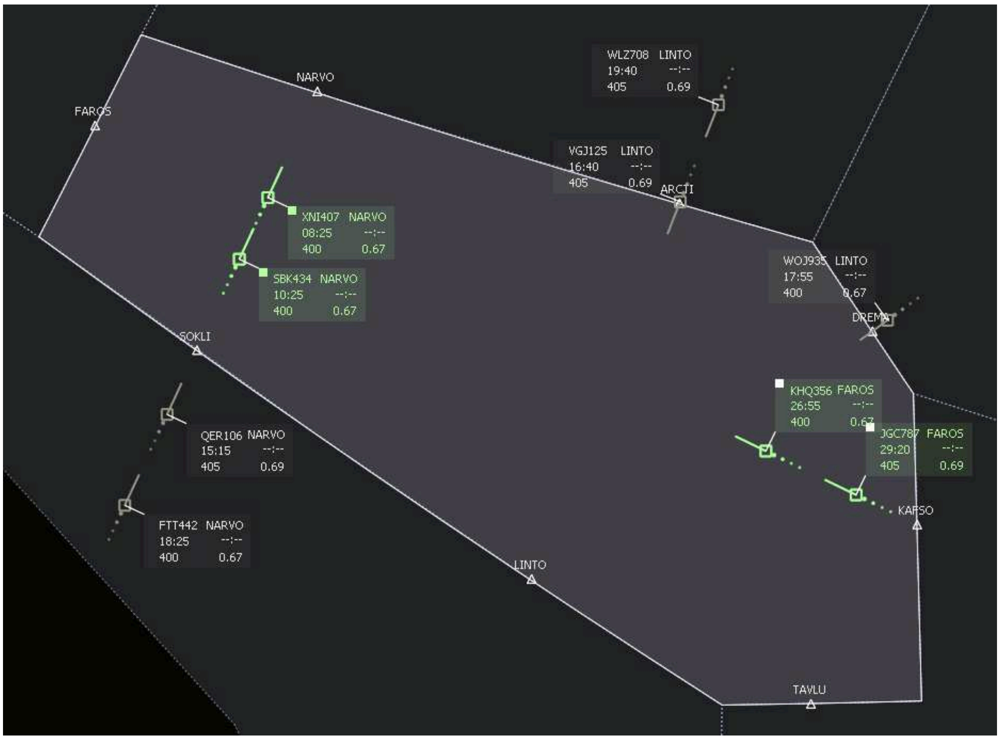


Figure B-2: Structured traffic with low perturbations scenario

could be selected anywhere within the airspace sector and two new solution spaces for each trajectory segment will appear. If two waypoints are created within the airspace sector then three solution spaces appear (one for each segment), if four waypoints then five solutions spaces appear, etcetera.

To check the worthiness and feasibility of performing this thesis experiment study, another previous experiment that was performed at TU Delft in the past with the Travel Space tool was analyzed and compared to see if the DD and TBX metrics had any commonalities; if the metrics gave exactly the same results there would little gained performing an experiment. The setup of this previous experiment had levels of low, medium, and high perturbation accompanied with unstructured or structured air traffic scenarios; six scenario cases in total. Twelve subjects consisting of four engineering Doctor of Philosophy students, four Luchtverkeersleiding Nederland - *Air Traffic Control the Netherlands* (LVNL) ATCOs, and four Nationaal Lucht- en Ruimtevaartlaboratorium - *National Aerospace Laboratory* (NLR) domain experts each testing all six scenarios. The scenario of structured air traffic with low perturbations is shown in Fig. B-2. Notice that there are no weather cells in the airspace sector and the air traffic is fairly structured.

The case of structured air traffic and medium perturbations is not shown as a static image of this scenario will not show a noticeable difference to the static image of the high perturbations scenario. The scenario of structured air traffic with high perturbations is shown in Fig. B-3. In this scenario there is a weather cell which has a simple shape of a circle and there are aircrafts that are by default heading towards it, causing the controller to take corrective actions to resolve the conflicts.

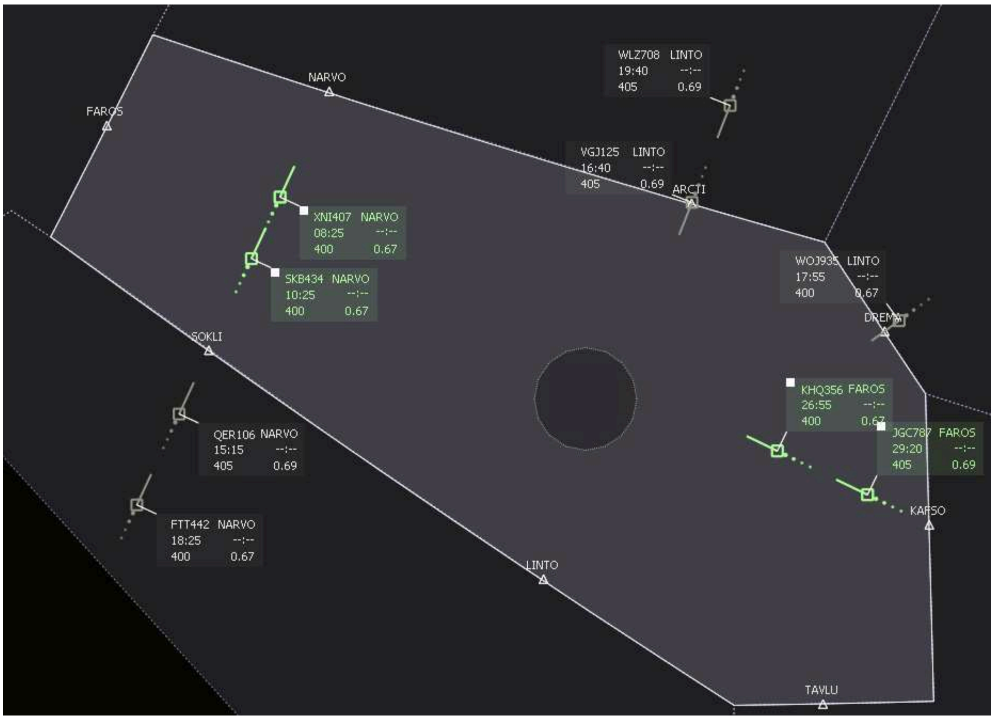


Figure B-3: Structured Traffic with large perturbations scenario

The scenario of unstructured traffic with low perturbations can be seen in Fig. B-4. Here the Travel Space tool can be seen in use with an aircraft selected showing its performance profile solution travel space with green and red areas. The corresponding conflicting aircraft is also highlighted in red making it visibly easy to identify. The scenario with unstructured traffic with medium perturbations is not shown for the same reason that it does not show a noticeable difference to the static image of the high perturbations scenario.

The scenario of the unstructured traffic with high perturbations can be seen in Fig. B-5. The selected aircraft shows the Travel Space tool being used and a weather cell in the airspace to add to the complexity.

Unfortunately, there were limitations with this previous experiment. The workload measurements of the experiment had low resolution with only three data points for each scenario run; beginning, middle, and end. These are not enough measurements to analyze and conclude if a complexity metric could correlate well to workload. A new experiment study was therefore justified and performed with a higher resolution of workload measurements (100 second scenario-time intervals).

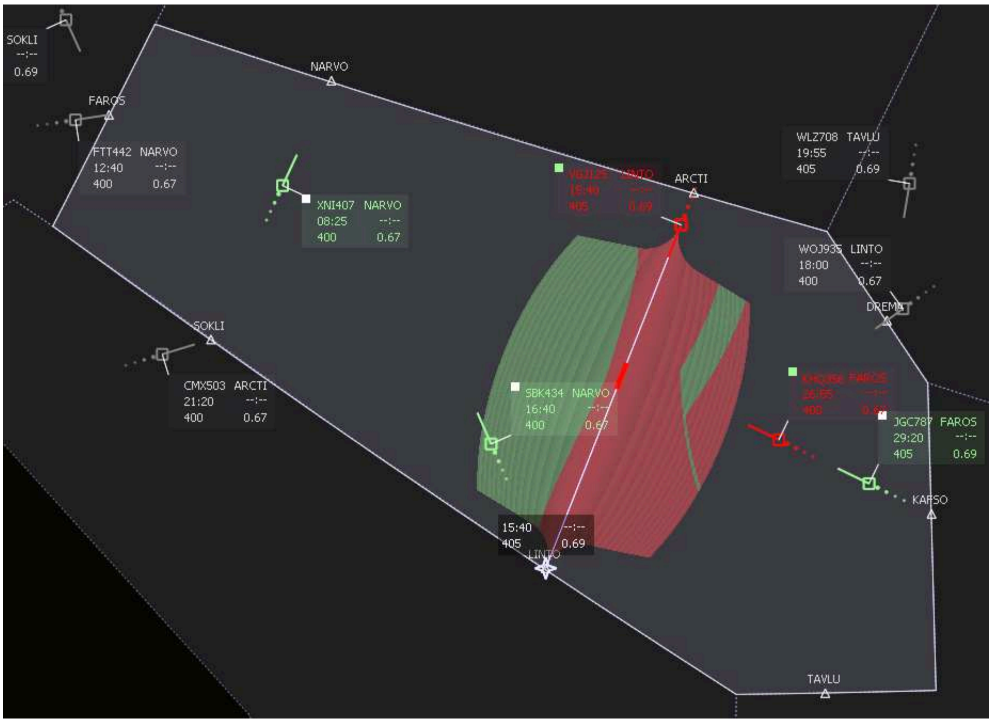


Figure B-4: Unstructured traffic with low perturbations scenario

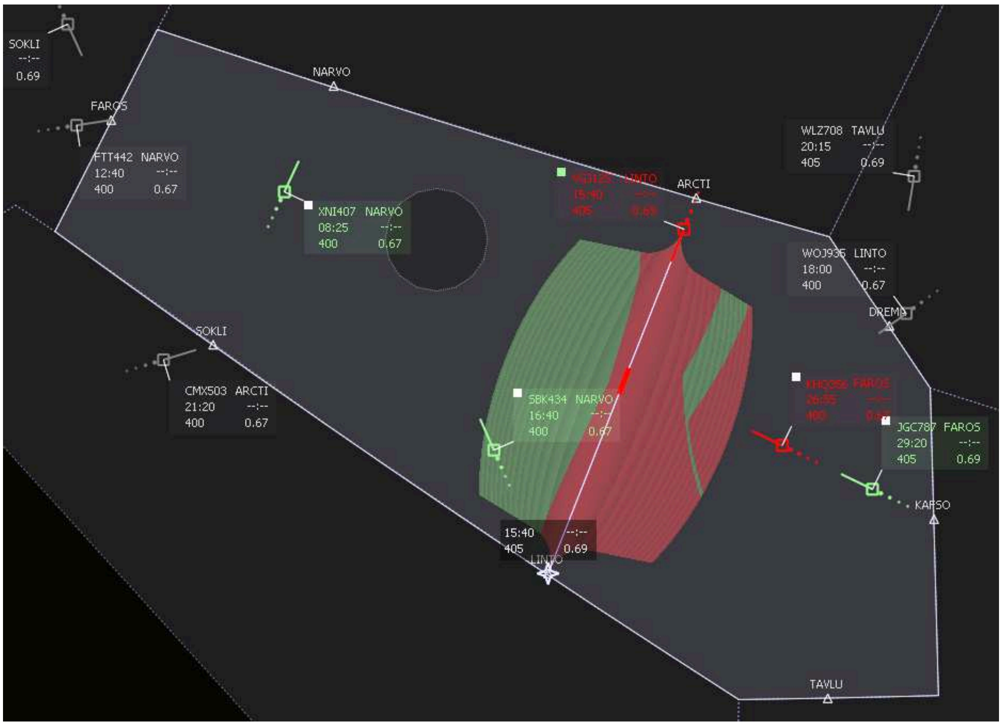


Figure B-5: Unstructured traffic with high perturbations scenario

Appendix C

Dynamic Density Metric

DD is the complexity or difficulty of a traffic situation and was first described as "the essential factors affecting conflict rate in both en route and terminal airspace" in a Radio Technical Commission for Aeronautics report. It is a collective effect of all factors, or variables, that contribute to the sector level ATC complexity or difficulty at any given time. It can also be described as a ATC taskload which is a function of the number of aircraft and the complexity of traffic patterns in a volume of airspace. (Kopardekar & Magyarits, 2002) However, this metric is state-based and not originally intended for 4D trajectory operations. Using DD a static snapshot of a given traffic scenario can be analyzed, but the dynamics of the system (given time-based intent) are not captured at all which trajectory management would make use of.

There have been multiple organizational research efforts to develop and validate several proposed DD metrics. The FAA developed a living document that provides a means for organizations to effectively use resources and eliminate duplication of effort. The organizations included William J. Hughes Technical Center (WJHTC), Titan Systems, National Aeronautics and Space Administration (NASA) Ames Research Center (ARC), Metron Aviation, and Mitre CAASD. The results have been four sets of dynamic density metrics that have been developed; WJHTC/Titan Metric (with 10 variables), NASA Metric 1 (with 16 variables), NASA Metric 2 (with 9 variables), and Wyndemere/Metron Metric (with 10 variables).

With the previous experiments (Abdul Rahman, Mulder, & Paassen, 2010) using DD it has been shown that NASA Metric 1 is the most promising of the DD metrics for the experimental conditions used for the TU Delft Travel Space Tool. The current form of the tool does not take into account altitude changes and assumes all aircraft are flying at the same flight level. Only the NASA Metric 1 has been considered in this preliminary study and is explained in more detail in the following paragraphs.

Table C-1: Valid NASA DD Metric 1 Elements for Travel Space

Element	Renamed	Description
C_1	Flow	Number of Aircraft Flow
C_5	HOR1	Inverse Mean Weighted Horizontal Separation Distance
C_7	HOR2	Inverse of the Average Minimum Horizontal Separation Distance
C_9	HOR3	Inverse of the Minimum Horizontal Separation
C_{11}	TTG1	Time-To-Go Fraction of Aircraft
C_{12}	TTG2	Time-To-Go Inverse of Average Minimum
C_{13}	TTG3	Time-To-Go Inverse of Smallest
C_{14}	SPD1	Variance of Groundspeed
C_{15}	SPD2	Groundspeed Ratio of Standard Deviation to Mean
C_{16}	MWN	Mean Conflict Resolution Difficulty

C-1 NASA Metric 1

The original metric consists of sixteen complexity elements that are summed: (C_1) ratio of number of aircraft, (C_2) ratio of aircraft count (AC) of climbing aircraft, (C_3) ratio of AC of cruising aircraft, (C_4) ratio of AC of descending aircraft, (C_5) inverse mean weighted horizontal separation distance, (C_6) mean weighted vertical separation, (C_7) inverse of the average minimum horizontal separation between aircraft pairs, (C_8) inverse of the average minimum vertical separation between aircraft pairs, (C_9) horizontal separation of aircraft within an altitude band, (C_{10}) vertical separation of aircraft in close horizontal proximity, (C_{11}) number of aircraft pairs with positive time-to-go, (C_{12}) average time-to-go, (C_{13}) smallest time-to-go, (C_{14}) variance of speed, (C_{15}) ratio of standard deviation of speed to average speed, and (C_{16}) conflict resolution difficulty based on crossing angle. (Chatterji & Sridhar, 2001)

In the Travel Space experiments it is assumed that all aircraft are at the same flight level. Therefore, left out are the ratios of (C_2) climbing, (C_3) cruising, and (C_4) descending aircraft and the elements with parts that consider vertical differences are equaled to zero or become no longer valid (C_6)(C_8)(C_{10}) for the experiment. This reduces the complexity elements to ten for the Travel Space which an overview can be seen in Table C-1. The modified DD NASA Metric 1 equation for the TU Delft Travel Space tool is given below.

$$
\begin{aligned}
DD_{TS} = a * Flow + b * HOR1 + c * HOR2 + d * HOR3 + e * TTG1 \\
+ f * TTG2 + g * TTG3 + h * SPD1 + i * SPD2 + j * MWN
\end{aligned}
\tag{C-1}
$$

Where the coefficients (a, b, c,...) are determined afterwards using linear regressions to fit the DD to a specific sector and are valid only for that sector. The following sections go into detail about the modified elements of Eq. (C-1).

C-1-1 Flow Complexity - Number of Aircraft

The number of aircraft in the sector is the simplest measure of flow complexity. This represents how many aircraft the ATCO will have to keep track of at any given time but does not fully capture the intricacies of complexity.

$$C_1 = Flow = \frac{N}{N_{max}} \tag{C-2}$$

Where N is the total number of aircraft within the sector at any instant of time. The term N_{max} is the historical, or the acceptable, maximum number of aircraft in the sector. There has been no modification to this element of the DD as it is straightforward for the Travel Space experiment and is now referred to as $Flow$. The value of $Flow$ will generally range from zero to one showing the capacity of the sector.

C-1-2 Horizontal Proximity Metric 1 - Inverse Mean Weighted Separation Distance

The inverse of the mean weighted horizontal separation between aircraft pairs is one of several proximity measures. The reasoning for using the inverse is that decreasing mean distances results from reduced separation between neighboring aircraft. Below are the original equations.

$$C_5 = \frac{N}{\sum_{1 \le i \le N} \left(\frac{\sum_{1 \le j \le N} W_{ij} d_{ij}}{\sum_{1 \le j \le N} W_{ij}} \right)} \tag{C-3}$$

$$W_{ij} = \begin{cases} i \neq j, (d_{ij}^2 + S_h^2 h_{ij}^2)^{-1} \\ i = j, 0 \end{cases} \tag{C-4}$$

C_5 is the inverse of the mean of the weighted mean horizontal separation distances. Where N is the number of aircraft within the sector, d_{ij} is the horizontal separation distance between the two aircraft i and j, W_{ij} is the associated weighting factor, h_{ij} is the vertical separation distance between the two aircraft i and j, and S_h is the scaling factor for making the altitude separation distance comparable to the horizontal separation distance. For altitudes in excess of 29,000 feet above mean sea level, the horizontal separation minimum is five nautical miles and the vertical separation is 2000 feet which gives $S_h = 0.0025 nm/ft$. The purpose of weighting is to reduce the contribution of aircraft that are farther away horizontally and vertically from the i^{th} aircraft and the neighboring aircraft with a bias towards neighboring aircraft.

In the Travel Space experiment all aircraft are at the same flight level so $h_{ij} = 0$. Below are the modified equations of Eqs. (C-3) and (C-4) for the Travel Space tool which is now called W_{ijTS} and $HOR1$.

$$W_{ijTS} = \begin{cases} i \neq j, (d_{ij}^2 + \cancelto{0}{S_h^2 h_{ij}^2})^{-1} = d_{ij}^{-2} \\ i = j, 0 \end{cases} \tag{C-5}$$

$$HOR1 = \frac{N}{\sum\limits_{1\le i\le N}\left(\frac{\sum\limits_{1\le j\le N} W_{ijTS}d_{ij}}{\sum\limits_{1\le j\le N} W_{ijTS}}\right)} = \frac{N}{\sum\limits_{1\le i\le N}\left(\frac{\sum\limits_{1\le j\le N}(\frac{1}{d_{ij}^2})d_{ij}}{\sum\limits_{1\le j\le N}(\frac{1}{d_{ij}^2})}\right)} = \frac{N}{\sum\limits_{1\le i\le N}\left(\frac{\sum\limits_{1\le j\le N} d_{ij}^{-1}}{\sum\limits_{1\le j\le N} d_{ij}^{-2}}\right)} \tag{C-6}$$

Notice that the scaling factor, S_h, of Eq. (C-4) goes away and W_{ij} reduces to $d_i^{-2}j$. Plugging this new W_{ij}, Eq. (C-5), into C_5 results to the new $HOR1$, Eq. (C-6). As the mean weighted separation distances decrease this will cause the value of $HOR1$ to increase, adding to the complexity contribution to the complete DD equation.

C-1-3 Horizontal Proximity Measure 2 - Inverse Average Minimum Separation Distance

The inverse average minimum horizontal separation between aircraft pairs is the second horizontal proximity measure. This measure is based on the average minimum separation between aircraft that would give more workload to the controller if on average the minimum separation are decreasing. Below are the original equations.

$$C_7 = \frac{\sum\limits_{1\le i\le N}[j \in J_i]}{\sum\limits_{1\le i\le N}\min\limits_{j\in J_i}\{d_{ij}\}} \tag{C-7}$$

$$J_i = \{j|h_i - \Delta h/2 \le h_{ij} \le h_i + \Delta h/2;\;\; j \ne i\} \tag{C-8}$$

C_7 is defined as the inverse of the average minimum horizontal separation between aircraft pairs. Where N is the number of aircraft within the sector, J_i is a set of aircraft that are within a Δh vertical neighborhood about the aircraft i, h_i is the altitude of aircraft i, h_{ij} is the vertical separation between the i and j aircraft, and d_{ij} is the lateral distance between the i and j aircraft within the Δh altitude band. The numerator counts the number of aircraft for which at least one other aircraft is found within its altitude neighborhood. Notice that the numerical value never exceeds N, which is the case when every aircraft has one or more aircraft in its altitude neighborhood.

Below is the modified equation of Eq. (C-7) for the Travel Space tool which is now called $HOR2$.

$$HOR2 = \frac{N}{\sum\limits_{1\le i\le N}\min\{d_{ij}\}} \tag{C-9}$$

Notice that the numerical value of the numerator of $HOR2$ becomes N because for J_i every aircraft are all at the same flight level for the Travel Space experiment. As the average minimum distances decrease, the value of $HOR2$ will increase, adding to the complexity contribution of the complete DD equation.

C-1-4 Horizontal Proximity Measure 3 - Inverse of Minimum Horizontal Separation

The inverse of minimum horizontal separation in the same vertical neighborhood is the final horizontal proximity measure. This measure is based on the minimum separation for a pair of aircraft within the group. This close separation between a pair of aircraft would cause the controller to focus the attention on this pair of aircraft because of the possibility of separation violation. Below is the original equation.

$$C_9 = \frac{1}{\min\limits_{1 \leq i \leq N} \left\{ \min\limits_{j \in J_i} \{d_{ij}\} \right\}} \tag{C-10}$$

C_9 is defined as the inverse of the minimum horizontal separation in the same vertical neighborhood. Where N is the number of aircraft within the sector, d_{ij} is the lateral distance between aircraft i and aircraft j, and for J_i refer to Eq. (C-8).

For the Travel Space experiment all aircraft are on the same flight level. Below is the modified equation of Eq. (C-10) for the Travel Space tool which is now called $HOR3$.

$$HOR3 = \frac{1}{\min\limits_{1 \leq i \leq N} \{d_{ij}\}} \tag{C-11}$$

Notice that the $min_{j \in J_i}$ goes away as they are all considered in the same vertical neighborhood and the equation simply takes the minimum distance between all the aircraft pairs. As the smallest distance between an aircraft pair out of the entire group decreases, the value of $HOR3$ will increase, contributing more to the complete DD equation.

C-1-5 Time-To-Go to Conflict Measure 1 - Fraction of Aircraft

The fraction of aircraft with time-to-go to conflict less than $\Delta t = 600s$ is the first of these kinds of measures that consider the urgency of conflict resolution. For this state-based approach the number and frequency of the sector controller's actions for conflict resolution can be based on the relative position, heading, and speed of aircraft pairs. These three qualities together determine if the aircraft pair would violate the separation minimum in the future.

The rate of change of distance between a pair of aircraft can be used to determine if the pair is moving toward each other or away from each other. In the cases where the aircraft are moving toward each other the pairs that are in close proximity with each other are more important compared to those aircraft pairs that are farther away. The closer pairs require immediate conflict resolution and the further away pairs can be postponed. Below is the original equation for the range rate, $\dot{d_{ij}}$.

$$\dot{d_{ij}} = \frac{(d_{xij}V_{xij} + d_{yij}V_{yij} + d_{hij}V_{hij})}{d_{ij}} \tag{C-12}$$

Where d_{ij} is the distance between the i and j aircraft pair and the terms d_{xij}, d_{yij}, and d_{hij} are the distance coordinates and V_{xij}, V_{yij}, and V_{hij} are the relative velocity components with respect to the body frame, attached to the i aircraft, measured along the inertial frame axes. The time-to-go to conflict t_{ij} can be determined in terms of the range rate given by the Eq. (C-12) as,

$$t_{ij} = -\frac{d_{ij}}{\dot{d}_{ij}} \tag{C-13}$$

From this equation, Eq. (C-13), it can be seen that time-to-go is positive if the closing rate is negative which indicates that the pair of aircraft are converging. From this equation several measures can be developed. The threshold time when conflict resolution becomes urgent is defined as Δt and T_i is the set of neighboring aircraft with time-to-go less than or equal to Δt seen below.

$$T_i = \{j | 0 \leq t_{ij} \leq \Delta t; \;\; j \neq i\} \tag{C-14}$$

Increasing numbers of such pairs can results in the monitoring workload to increase. Below is the flow complexity measure, C_{11}, derived from this increase. C_{11} is based on the number of positive time-to-go less than or equal to Δt.

$$C_{11} = \frac{\sum\limits_{1 \leq i \leq N} \; \sum\limits_{j \in T_i, J_i} 1}{2N} \tag{C-15}$$

N is the number of aircraft within the sector and the factor of 2 is used in the denominator because the pairs i an j are counted twice. J_i is a set of aircraft indices such that their corresponding altitudes are within the vertical neighborhood of aircraft i, refer to Eq. (C-8). The term T_i is the indices with time-to-go less than the threshold value, refer to Eq. (C-14). The time-to-go threshold Δt is set to 600s, t_{ij} is the time-to-go, d_{ij} is the distance between the i and j aircraft pair, $\dot{d}_{ij}$ is the range rate.

For the Travel Space experiment the height components d_{hij} and V_{hij} will go away from Eq. (C-12) as the altitude for all aircraft are at the same flight level. The modified range rate equation is seen below.

$$\dot{d}_{ijTS} = \frac{(d_{xij}V_{xij} + d_{yij}V_{yij} + \cancelto{0}{d_{hij}V_{hij}})}{d_{ij}} = \frac{(d_{xij}V_{xij} + d_{yij}V_{yij})}{d_{ij}} \tag{C-16}$$

Furthermore, because the aircraft are at the same flight level each aircraft are part of J_i. Therefore, J_i is simply left out and the modified equation of Eq. (C-15) for the Travel Space tool is written below which is now called $TTG1$.

$$TTG1 = \frac{\sum\limits_{1 \leq i \leq N} \; \sum\limits_{j \in T_i} 1}{2N} \tag{C-17}$$

Notice as the number of aircraft pairs increase the value of $TTG1$ increases, adding to the complexity of DD but will never exceed a value of one half.

C-1-6 Time-To-Go to Conflict Measure 2 - Inverse Minimum

The inverse minimum time-to-go conflict with less than $\Delta t = 600s$ is another measure that considers the urgency of conflict resolution. This measurement is based on the average time-to-go value on a given set of aircraft pairs with less than 600 seconds which indicates the time the controller has for resolving conflict in general. Below is the original equation.

$$C_{12} = \frac{\sum\limits_{1 \leq i \leq N} [j \in T_{i,J_i}]}{\sum\limits_{1 \leq i \leq N} \min\limits_{j \in T_i} \{t_{ij}\}} \tag{C-18}$$

N is the number of aircraft within the sector. The term T_i is the set of neighboring aircraft with time-to-go less than or equal to 600 seconds and J_i is the set of aircraft that are within a Δh vertical neighborhood about the aircraft i, refer to Eqs. (C-8) and (C-14) respectively. The time-to-go to conflict is t_{ij}, refer to Eqs. (C-12) and (C-13). The numerator counts the number of aircraft that have at least one other aircraft in the set defined by Eq. (C-14). The index of summation i is bound by 1 and N therefore, the maximum value of the numerator is N. For the denominator the minimum time-to-go between the i aircraft and all members in the set T_i is computed for every aircraft. The resulting values are summed up to obtain the value for the denominator.

For the Travel Space experiment all aircraft are at the same flight level so there is no need to mention J_i. Furthermore, the modified range rate $\dot{d}_{ijTS}$ is used instead as seen in Eq. (C-16) when calculating t_{ij}. The modified equation of Eq. (C-18) for the Travel Space tool is written below which is now called $TTG2$.

$$TTG2 = \frac{\sum\limits_{1 \leq i \leq N} [j \in T_i]}{\sum\limits_{1 \leq i \leq N} \min\limits_{j \in T_i} \{t_{ij}\}} \tag{C-19}$$

The value of $TTG2$ represents the inverse average minimum time to conflict for aircraft pairs with a time-to-go less than 600 seconds. If there are no conflicts within 600 seconds, the value of the measure is zero. When the average time to conflict between aircraft pairs becomes smaller (i.e., conflict is more imminent), this measure increases the overall DD score.

C-1-7 Time-To-Go to Conflict Measure 3 - Inverse of Smallest

The inverse of smallest time-to-go conflict for aircraft pairs with time-to-go to conflict less than $\Delta t = 600s$ is similar to Eq. (C-10). This measurement is based on the logic that the workload experienced by the controller will be higher, if the time available for resolution of the immediate conflict is smaller. Below is the original equation.

$$C_{13} = \frac{1}{\min\limits_{1 \leq i \leq N} \left\{ \min\limits_{j \in T_i, J_i} \{t_{ij}\} \right\}} \tag{C-20}$$

N is the number of aircraft within the sector. The term T_i is the set of neighboring aircraft with time-to-go less than or equal to 600s and J_i is the set of aircraft that are within a Δh vertical neighborhood about the aircraft i, refer to Eqs. (C-8) and (C-14) respectively. The time-to-go is t_{ij}, refer to Eqs. (C-12) and (C-13). C_{13} is zero when none of the times-to-go are smaller than $\Delta t = 600s$.

For the Travel Space experiment all aircraft are at the same flight level so there is no need to mention J_i. Furthermore, the modified range rate $\dot{d}_{ijTS}$ is used instead as seen in Eq. (C-16) when calculating t_{ij}. The modified equation of Eq. (C-20) for the Travel Space tool is written below which is now called $TTG3$.

$$TTG3 = \frac{1}{\min\limits_{1 \le i \le N} \left\{ \min\limits_{j \in T_i} \{t_{ij}\} \right\}} \quad \text{(C-21)}$$

The value of $TTG3$ represents the inverse of the smallest time to conflict for aircraft pairs with a time-to-go less than 600 seconds. If there are no conflicts within 600 seconds, the value of the measure is zero. When the average time to conflict between aircraft pairs becomes smaller (i.e., conflict is more imminent), this measure increases the overall DD score.

C-1-8 Speed Measure 1 - Variance of Groundspeed

The variance of groundspeed is how far the values of aircraft groundspeed are spread out amongst each other. For DD it is thought that the variance of groundspeed is possibly a good measure of complexity. Average speed and average sector transit time have been used as measures of complexity by some researchers in the past but a study reviewing these measures for DD concluded that these may not be a good measure of complexity because the variability in the parameters within the sector was much lower than between sectors. The study also found that the mean airspeed was not significantly correlated to the behavioral response so it did not give a good measure of job difficulty. (Chatterji & Sridhar, 2001) The variance of groundspeed equation is given below.

$$C_{14} = \sigma_{vg}^2 = \frac{\sum\limits_{1 \le i \le N} (V_i - \bar{V})^2}{(N-1)} \quad \text{(C-22)}$$

$$V_i = \sqrt{V_{xi}^2 + V_{yi}^2 + V_{hi}^2} \quad \text{(C-23)}$$

$$\bar{V} = \frac{\sum\limits_{1 \le i \le N} V_i}{N} \quad \text{(C-24)}$$

In Eq. (C-22), σ_{vg}^2 is the variance of groundspeed. N is the number of aircraft within the sector. V_i is the groundspeed of aircraft i and $\bar{V}$ is the mean of groundspeed. The groundspeed and mean groundspeed are defined respectively in Eqs. (C-23) and (C-24). In the groundspeed equation velocity components are given in terms of x, y, and h for an i^{th} aircraft.

For the Travel Space experiment all aircraft are at the same flight level so there is no V_{hi}^2 component in the groundspeed equation. The modified equation of Eq. (C-23) for the Travel Space tool is given below and Eq. (C-22) is now called $SPD1$.

$$V_i = \sqrt{V_{xi}^2 + V_{yi}^2 + \cancelto{0}{V_{hi}^2}} = \sqrt{V_{xi}^2 + V_{yi}^2} \tag{C-25}$$

$$SPD1 = \sigma_{vg}^2 = \frac{\sum\limits_{1 \le i \le N} (V_i - \bar{V})^2}{(N-1)} \tag{C-26}$$

The value of $SPD1$ represents how far the values of aircraft groundspeed are spread out. The value can have a wide range depending on the amount of aircraft in the sector, affecting the average speed value and the speed differences. This measure contributes to the complete DD equation and can have a large value although each DD measure has a coefficient fitted to the sector which is determined after many runs and the data linearly regressed to workload.

C-1-9 Speed Measure 2 - Groundspeed Ratio of Standard Deviation to Mean

The ratio of standard deviation of ground speed to mean of ground speed is known as the contrast ratio. The measure models the heuristic that higher variance does not increase workload if the average groundspeed is low. The measure is based on the variance and the mean of the groundspeed which is given below.

$$C_{15} = \frac{\sigma_{vg}}{\bar{V}} \tag{C-27}$$

$$\sigma_{vg} = \sqrt{\frac{\sum\limits_{1 \le i \le N} (V_i - \bar{V})^2}{(N-1)}} \tag{C-28}$$

In Eq. (C-27), σ_{vg} is the standard deviation and $\bar{V}$ is the mean groundspeed. Refer to Eq. (C-24) for the definition of mean groundspeed. In Eq. (C-28), N is the number of aircraft within the sector and V_i is the groundspeed of the i^{th} aircraft.

For the Travel Space experiment all aircraft are at the same flight level so Eq. (C-25) is used for V_i and Eq. (C-27) is now called $SPD2$ as seen below.

$$SPD2 = \frac{\sigma_{vg}}{\bar{V}} \tag{C-29}$$

The value of $SPD2$ becomes very small, a fraction of the value one, because of the square root and divided by the mean groundspeed. This measure contributes to the complete DD equation and although the value is small, each DD measure has a coefficient fitted to the sector which is determined after many runs and the data linearly regressed to workload.

C-1-10 Mean Conflict Resolution Difficulty

The mean conflict resolution difficulty is based on crossing angle of two aircraft. Small crossing angle conflicts are the most complex and 90 degree crossing angle conflicts are the least complex. With head-on conflicts to be high complexity because of high closing rates. Shallow crossing angle conflicts are harder to detect but easier to resolve because the controller has more time for path corrections while head-on conflicts are easy to detect but harder to resolve. A study has found shallow crossing angle resolutions have to be initiated earlier compared to large crossing angles. A normalization factor, n_f, of $3208.2s$ (53.47 min) has been used based on a graph of normalized time of resolution initiation as a function of cross angle. (Chatterji & Sridhar, 2001)

The crossing angle ξ_{ij} for the i and j pair of aircraft is given by $\xi_{ij} = min(|\chi_{ij}|, 2\pi - |\chi_{ij}|)$ where χ_{ij} is the relative heading angle of the converging pair of aircraft. The time-to-go t_{ij} as defined in Eq. (C-13) and using Eq. (C-16) is less than the time-to-go threshold Δt in the set T_i defined in Eq. (C-14). The heading angle for i aircraft is defined below.

$$\chi_i = tan^{-1}\left(\frac{V_{yi}}{V_{xi}}\right) \tag{C-30}$$

Where V_{xi} and V_{yi} are the velocity vector components on the horizontal plane.

The level of resolution difficulty as a function of crossing angle is obtained from the mentioned graph which is available as a lookup table. Each of the converging pair of aircraft can be combined to define the overall complexity measure associated with conflict resolution. The mean conflict resolution difficulty is given below.

$$C_{16} = \frac{\displaystyle\sum_{1\leq i\leq N}\sum_{j\in T_i, J_i} \varpi_{\xi_{ij}} n_f}{2N} \tag{C-31}$$

N is the number of aircraft in the sector and T_i is the set of neighboring aircraft with time-to-go less than or equal to $\Delta t = 900s$. The term J_i is a set of aircraft in a vertical neighborhood that is defined in Eq. (C-8). The level of resolution difficulty is given as $\varpi_{\xi_{ij}}$ and the normalization factor is $n_f = 3208.2s$. A factor of 2 has been included in the denominator to account for the fact that i and j aircraft pair is counted twice.

For Travel Space experiments the aircraft are all at the same flight level. Therefore, J_i is simply left out as all aircraft are part of the set. The modified equation of Eq. (C-31) for the Travel Space tool is written below which is now known as MWN.

$$MWN = \frac{\displaystyle\sum_{1\leq i\leq N}\sum_{j\in T_i} \varpi_{\xi_{ij}} n_f}{2N} \tag{C-32}$$

The value of MWN represents the mean conflict resolution difficulty and the level of the resolution difficulty, ($\varpi_{\xi_{ij}}$) is determined by the cross angle and a lookup table. When there are head-on or small crossing angle conflicts between a pair or pairs of aircraft, this measure increases the overall DD score.

Appendix D

Trajectory-based Complexity Metric

TBX is a modified aircraft count that unlike DD can be computed and communicated easily to predict in real-time sector complexity for a trajectory-based operations (TBO) environment. NextGen will incorporate TBO, which is a new concept that will need to be investigated for a new approach to measure traffic complexity in order to help design new trajectory-based systems. Previous metrics of traffic complexity such as DD are state-based. TBX is a recent metric developed and evaluated in the Airspace Operations Laboratory (AOL) at NASA ARC during human-in-the-loop studies of trajectory-based concepts since 2009. This metric considers more dynamic factors such as weather, aircraft equipage, predicted separation violations, as well as static factors such as sector size. It is a better predictor of workload than aircraft count and works well for complexity management in trajectory-based operations that can readily adjust to future operational concepts. Other metrics of complexity in the past such as DD had to be computed off-line and matched to previously collected data rather than in real-time based on current and predicted data. Resulting metrics that provided good statistical fit to the historical data were often ill-suited to characterize the uncertainties associated with traffic predictions. (Prevot & Lee, 2011)

In developing TBX the pitfalls of prior approaches to complexity calculations were avoided while keeping some of the features that work well for predicted complexity calculations that human operators could use for traffic flow management (TFM) purposes. The metrics were design to: be comprehensible to human operators, have weightings that would be easy to modify for new equipage and operational environments, be stable over the prediction time horizon, use TBO to create more stable trajectory predictions over a 1-2 hour time horizon, and combine different complexity factors into a single modified aircraft count value that human operators could use in conjunction with a Monitor Alert Parameter (MAP) value. To support real-time TFM, TBX was designed for real-time computations of complexity estimates based upon predictions of 4D trajectories for weather and aircraft.

For future experiments using the TU Delft Travel Space tool the TBX metric looks promising to be one of the best fit metrics to measure complexity as it was specifically designed for trajectory-based concepts. Before the TBX can be computed it requires prior determinations of defining the nominal conditions, where aircraft count is a good predictor of controller workload, and defining adjustments to the nominal conditions that capture the difference in complexity between the current conditions and the nominal conditions. An airspace sector's predicted aircraft count is used to determine the maximum traffic threshold that a controller can handle in that sector; this maximum threshold value is called MAP and is set based on average sector flight time.

The adjustments are ratios that describe the relationship between the current value and the nominal value. There are two types of adjustments, primary complexity adjustments and secondary complexity adjustments. Primary complexity adjustments are factors that can cause a similar effect to the complexity as the aircraft count, can offset the aircraft count, or have a detrimental impact. Theses factors are expected to have a multiplicative effect on the aircraft count. Weather or sector size are an example of dominant items that have primary adjustments. Secondary adjustments are factors that have a secondary/smaller impact on complexity which can be specific to a given operational environment. Their impact are less dependent on the overall aircraft count such as number of conflicts or transitioning aircraft.(Prevot & Lee, 2011)

In Eq. (D-1) the general form of TBX is shown. The TBX value of a sector s at time t equals the sum of its predicted aircraft count (ac) multiplied by the product of the primary adjustments (px) and the nominal aircraft count multiplied by the weighted sum of the secondary adjustments (sx) minus one. TBX and its adjustments have the value of 1 under nominal conditions which satisfies Eq. (D-2). All adjustments px_i and sx_j are limited to a range of 0 to 2 to prevent one factor from dominating the overall TBX value. (Prevot & Lee, 2011)

$$tbx(s,t) = ac(s,t) * \prod_{i=0}^{n} px_i(s,t) + ac_{nom}(s) * \left(\frac{\sum_{j=0}^{m}(w_j * sx_j(s,t))}{\sum_{j=0}^{m}(w_j)} - 1 \right) \tag{D-1}$$

$$tbx_{nom}(s,t) = ac(s,t) \tag{D-2}$$

where

s	is the sector of interest
t	is the time for which the complexity is computed
$tbx(s,t)$	is the predicted trajectory-based complexity of sector s at time t
$ac(s,t)$	is the predicted aircraft count of sector s at time t
$px_i(s,t)$	is the adjustment for primary complexity item i in sector s at time t
$ac_{nom}(s)$	is the nominal aircraft count in sector s
w_j	is the weight of secondary complexity item j
$sx_j(s,t)$	is the adjustment for secondary complexity item j in sector s at time t

In this structure of Eq. (D-1), MAP values can be based on nominal conditions for a given sector and then compare aircraft count and TBX value relative to the MAP to assess the peak that exceeds the MAP. This representation has the same format as aircraft count making the interpretation of TBX simple.

D-1 TBX Calculation

For the TU Delft Travel Space tool the general Eq. (D-1) will have the complexity factors of aircraft count, sector area, aircraft predicted to penetrate weather, and predicted loss of separation (LOS) events. Below is the modified equation for the Travel Space tool.

$$tbx(s,t)_{TravSP} = ac(s,t) * px_{sa}(s,t) * px_{wx}(s,t) + ac_{nom}(s) * \left(\frac{sx_{PrLos}(s,t)}{1} - 1\right) \quad \text{(D-3)}$$

Where s is the sector of interest and t is the time for which the complexity is computed. For the previous Travel Space experiment performed $ac_{nom}(s)$, the nominal aircraft count in sector s, is defined to have 80% of the MAP value and set MAP = 10 throughout the complexity factors ($ac_{nom}(s) = 8$). In future experiments performed using the Travel Space tool the $ac_{nom}(s)$ and MAP values will be determined by the experimental setup and conditions. The following complexity factor sections explain the rest of the terms in Eq. (D-3) in detail.

D-1-1 Aircraft Count

The aircraft count is a simple way to show the amount of aircraft that the ATCO must keep track of and how saturated the airspace is to deal with. Although this does not necessarily capture all the intricate qualities of airspace complexity, it has a major influence on it.

$$ac(s,t) = N \quad \text{(D-4)}$$

Where N is the number of predicted aircraft within the sector s at time t. TBX is based upon the assumption that aircraft count is the best predictor of workload under nominal conditions. Notice that aircraft count is the dominant complexity factor as seen in Eq. (D-3).

D-1-2 Sector Area

The usable sector area can have an influence on complexity since larger usable sectors provide more airspace than smaller ones for maneuvering aircrafts, communication, and time to decide on new/modified trajectories.

$$px_{sa}(s,t) = \left(\frac{sa(s,t)}{sa(s,t) - sua_{wx}(s,t)}\right)^{0.15} \quad \text{(D-5)}$$

Where $px_{sa}(s,t)$ is a primary complexity adjustment for sector area in sector s at time t. The term $sa(s,t)$ is the sector area of the current sector s at time t. The acronym sua stands for special use airspace area of the sector which can include government zones, military zones, weather, etc. The term $sua_{wx}(s,t)$ is the special use airspace area of weather for the current sector s at time t. The Travel Space tool experiment only uses weather cells, if there were multiple kinds of special use airspace areas then these would be added with each other. (e.g., $sua_{gov}(s,t) + sua_{mil}(s,t) + sua_{wx}(s,t)$) The power of 0.15 has been chosen so that the $px_{sa}(s,t)$ value does not dominate the overall tbx_{TravSP} value. The power value can be tuned empirically with more studies by assessing the impact of sector size on the overall complexity value and comparing the complexity to the corresponding controller workload. Notice if there are no weather cells, $sua_{wx}(s,t) = 0$, then the value of px_{sa} becomes 1 and does not contribute to the tbx_{TravSP} value. As the value of $sua_{wx}(s,t)$ increases the px_{sa} value rises exponentially. In other words, the bigger the weather cell becomes the more complex it will be to reroute air traffic with the decreasing usable sector area.

D-1-3 Weather

The weather impacts complexity by restricting the available airspace, which causes the need for rerouting flights and often results in substantial increases in communication between pilots and controllers. Also, weather becomes less predictable the further out into the future the prediction is made. The number of aircraft predicted to penetrate weather within a selected sector is used in the TBX calculation. The predicted weather trajectory is compared against the predicted aircraft trajectory to determine whether these two intersect within a sector.

$$px_{wx}(s,t) = 1 + 2 * \left(\frac{ac_{wx}(s,t)}{MAP}\right) \tag{D-6}$$

Where $px_{wx}(s,t)$ is a primary adjustment for the weather penetrations in sector s and at time t. The term $ac_{wx}(s,t)$ is the number of aircraft predicted to penetrate the weather of the current sector s at time t. The MAP value is set to 10 aircraft. Notice if half of a sectors MAP value will penetrate the weather, the complexity is doubled in comparison. If the maximum value of aircraft were to penetrate the weather, $ac_{wx}(s,t) = MAP$, it would triple the complexity.

D-1-4 Conflicts

The conflicts between aircraft add to complexity by causing controllers to assess the situation and at least requiring them to maneuver one of the conflicting aircraft. Often this includes coordinating with adjacent sectors, further increasing the workload. Trajectory automation is used to predict conflicts only reliably for up to 20 or 30 minutes because of uncertainties. (Prevot & Lee, 2011) With the Travel Space tool data collection setup we only consider one type of conflict related data because the current version of the tool only focuses on conflicts predicted within the sector and does not account for aircraft that have a conflict in a different sector. Below is the equation for the number of predicted LOS events.

$$sx_{PrLos}(s,t) = 1 + 2 * \left(\frac{PrLos(s,t)}{MAP} - PrLos_{nom}\right) \tag{D-7}$$

Where $sx_{PrLos}(s,t)$ is a secondary adjustment for the predicted number of LOS that occur in sector s at time t. The term $PrLos(s,t)$ is the number of predicted LOS events to occur in the sector s at time t, if no action is taken. The term $PrLos_{nom}(s,t)$ is the nominal value for predicted LOS to occur in a sector s at time t. We set the value of $PrLos_{nom}$ to 0.05 for our experiments which means the number of conflicts is about as much as 5% of the MAP value. The MAP value is set to 10 aircraft. Notice the limit of $sx_{PrLos} = 2.9$ when $PrLos = MAP$ and in the secondary adjustment portion of Eq. (D-3) it reduces to 1.9.

Appendix E

Justification for New Experiment Study

The use of the TBX metric created by NASA is new to TU Delft and has never been used before, so previous data from a past experiment (Abdul Rahman et al., 2010) at TU Delft that used DD were used for the initial investigation of the TBX metric. However, this previous experiment data had few workload rating measurement points (beginning, middle, and end of a scenario) to show any useful trends of how well the complexity metrics fit to what the user actually experienced in order to validate both metrics. Despite this limitation it was still worth comparing the two different complexity metrics to one another initially to see how much they have in common to each other. If it is the case that DD and TBX correlate completely well to each other then it would not prove fruitful to purse a new experiment. The results clearly showed this was not the case and in the following section the differences and trends are explained in detail.

E-1 Dynamic Density vs Trajectory-based Complexity

DD is state-based which was designed for todays ATM operations, not 4D trajectory management. This is a downside as it was not meant for predicting 4D trajectory intentions. Another downside to this metric is that it is calculated post processing using statistical regression analysis to determine weightings which are valid only for fitted sectors. These weightings are determined empirically after numerous runs of a certain sector in order to predict future complexity that fits with workload ratings.

TBX on the other hand was specifically designed for trajectory-base measurements and is computed in real-time. This is useful as it does not need to be bound to a specific sector such as with the DD. These qualities sounds promising to 4D trajectory management; however, it does not mean it will work well with the experiment conditions used in current TU Delft 4D experiments. (i.e., all aircraft at the same flight level, no aircraft equipage issue, etc.) Using

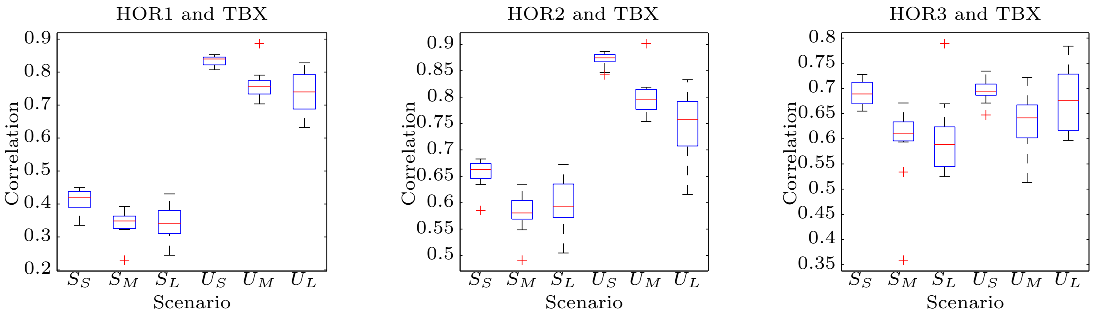


Figure E-1: Correlations of DD HORs with Full TBX

the past experimental data to compare the complexity metrics resulted in some interesting correlation trends between DD and TBX.

In the data there is not enough workload ratings data to determine good fit weighting values for the elements of the DD, so the raw values of each individual element was correlated to the full TBX value (Fig. J-1) as well as with the individual elements of the TBX. (Appendix A) In Fig. E-1 it can be seen that there is a clear trend with the three Horizontal Proximity Metric (HOR) elements. The scenarios 1 through 6 in each plot are structured traffic, low perturbations; structured traffic, medium perturbations; structured traffic, high perturbations; unstructured traffic, low perturbations; unstructured traffic, medium perturbations; and unstructured traffic, high perturbations respectively. The trend that is so striking is the decreasing correlation values between the increasing levels of perturbations in structured air traffic (scenario 1,2,3) and a significant correlation value jump to the set of unstructured air traffic with increasing levels of perturbations (scenario 4,5,6) but with the same decreasing trend as seen in HOR1, HOR2, and slightly with HOR3.

To further investigate the reason for this trend the TBX elements were separately correlated to the raw DD elements. In Figs. E-2, J-6 and J-7 in can be seen that the TBX aircraft count is the driving factor in this trend. The correlation of HOR1 and aircraft count in Fig. E-2 resembles the HOR1 with the full TBX in Fig. E-1. The same can be seen in HOR2 and aircraft count in Fig. J-6 with the HOR2 with the full TBX in Fig. J-1. As well as with HOR3 and aircraft count in Fig. J-7 resembles HOR3 with the full TBX in Fig. J-1.

When plotting HOR values with the TBX values of the scenarios with the strongest correlations in structured and unstructured traffic (scenario 1 and 4) a strong pattern can be seen as well. A sample of this ca be seen in Fig. E-3 where there is a jump each time in the HOR value around the 250s mark in scenario 1 which is observed across all research subjects. (Appendix B)

On further review of the playbacks of the scenario runs and the calculations of the HOR values it confirms the trends in the graphs and correlation figures. These trends are due to the calculations of the aircrafts average/minimum distances from each other which is affected when an aircraft enters or leaves the sector resulting in the value of the HOR going up or down. In Fig. E-4 the air traffic is structured and as such the aircraft are closely spaced together when heading the same direction as well as on average as a whole the average distances calculated are small. Looking at Fig. E-5 the air traffic is unstructured and results in aircraft spaced further apart because they are entering/exiting the sector at different points resulting

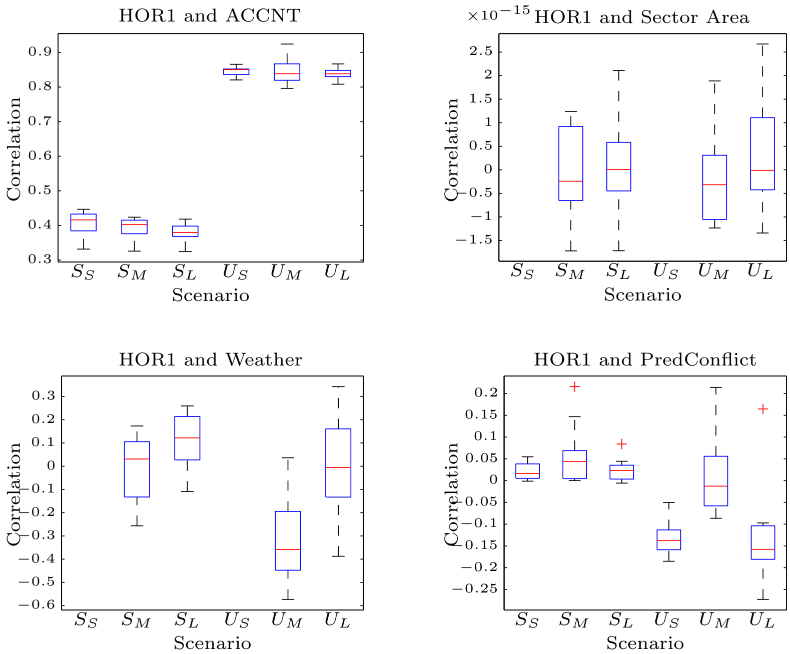


Figure E-2: Correlations of DD HOR1 with TBX elements

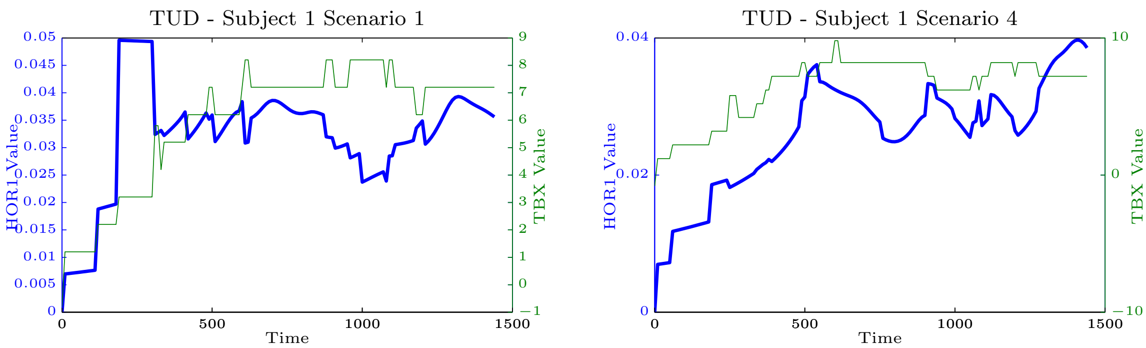


Figure E-3: HOR1 plotted against TBX

Figure E-4: Scenario 1 - structured traffic/low perturbations

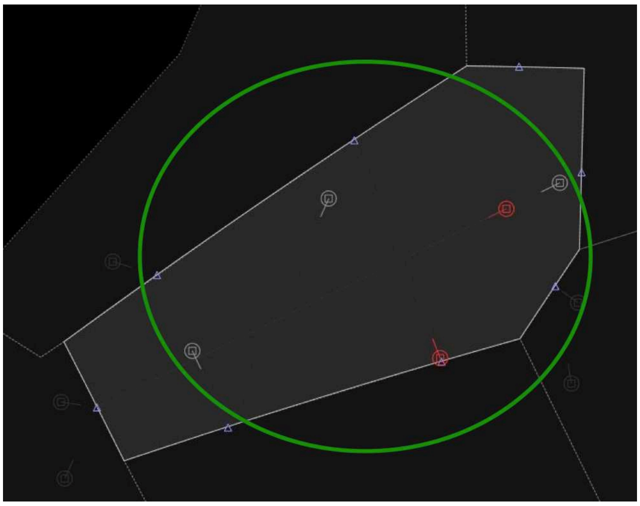


Figure E-5: Scenario 4 - unstructured traffic/low perturbations

in larger average distances calculated. The aircraft entering/exiting explains why the TBX value in general has a good correlation with the DD HOR values as the TBX value is mainly based on a modified aircraft count.

The results of this initial study between the two complexity metrics shows that they are fairly different from each other but it does not say anything about how well either fits to human workload. This gives reason to perform a new experiment that has high workload resolution data in order to finally compare the complexity metrics in a 4D trajectory-based environment and see which is the best fit. Because DD is state-based it is predicted that TBX will result in a better fit to workload in a 4D trajectory-based environment. It could be that there is not one best fit complexity metric; but instead a combination of the two in certain cases that as a whole fit well to workload.

Appendix F

Experiment Study Training and Scenario Runs

The research experiment study was performed on the 26th of March 2015 going till the 2nd of April 2015. The total experiment time for a participant was approximately 1.5-2 hours depending on how much training they needed to become familiar with the Travel Space tool. There were 16 participants in total (12 males, 4 females, average age of 25) who were all TU Delft aerospace master students with some knowledge of how ATM functions but no prior training in operational air traffic control. Following are examples of the training runs and experiment scenarios.

F-1 Training Runs

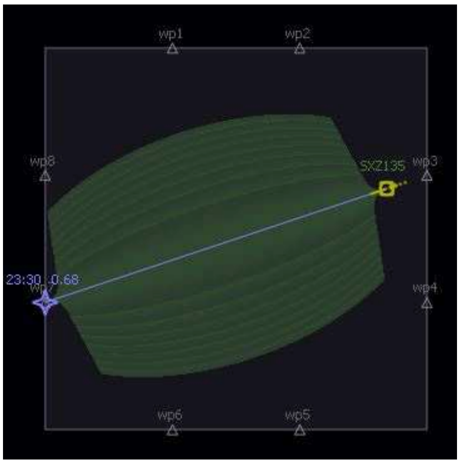
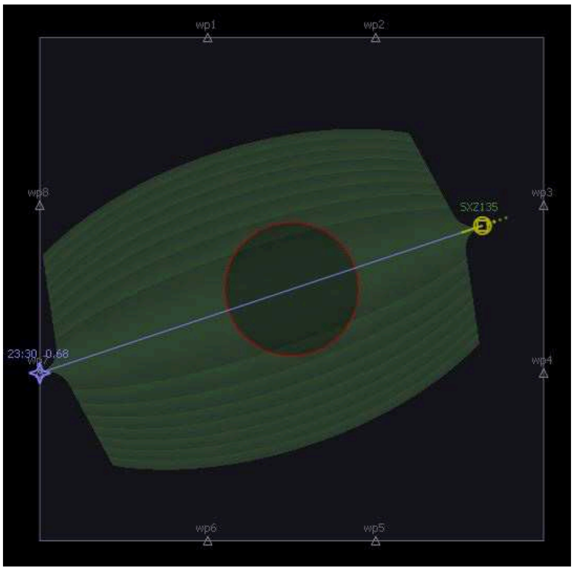
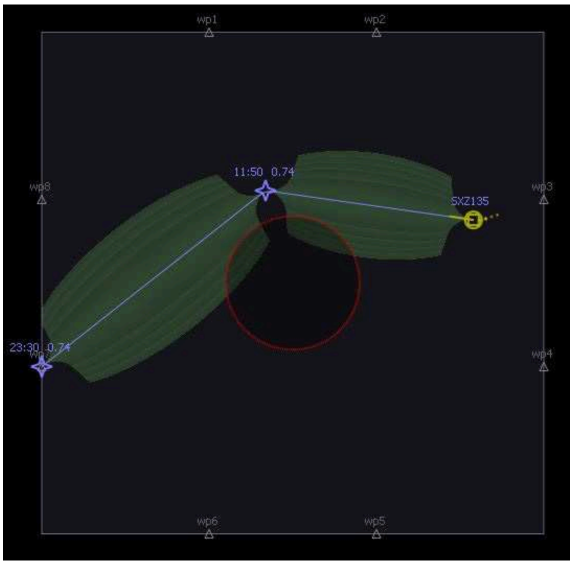


Figure F-1: Training 1 - Introduction to system and representations.

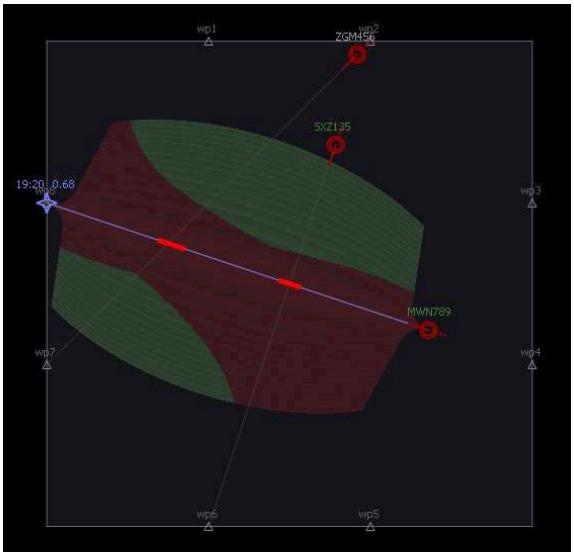


(a) Selecting aircraft with a restricted area in the sector.

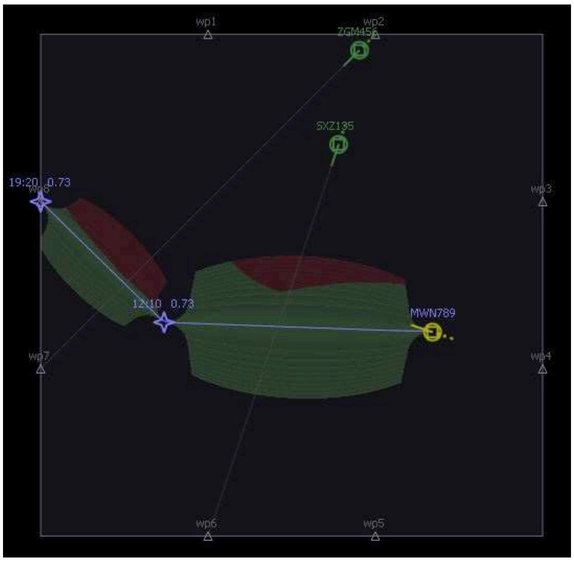


(b) Selecting a waypoint to reroute aircraft around restricted area.

Figure F-2: Training 2 - Introduction to restricted area and using waypoints.

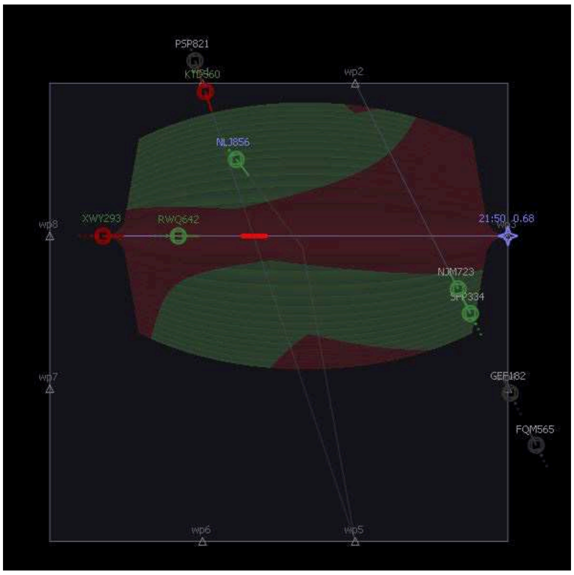


(a) Conflicts with multiple aircraft.

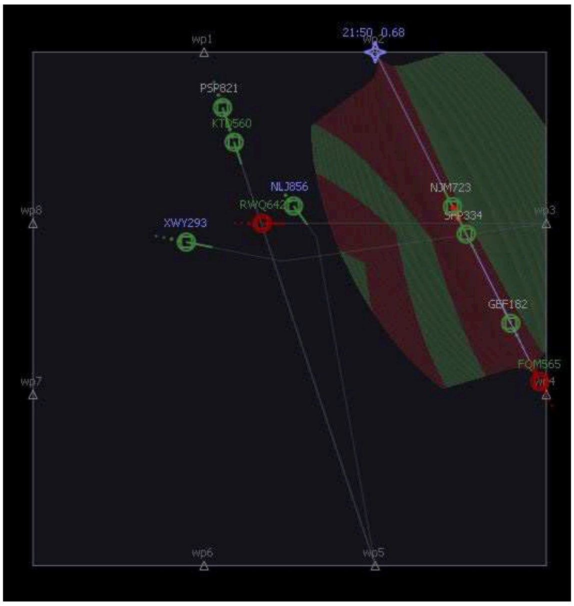


(b) Conflict resolution via a waypoint.

Figure F-3: Training 3 - Conflict resolution and small perturbation.

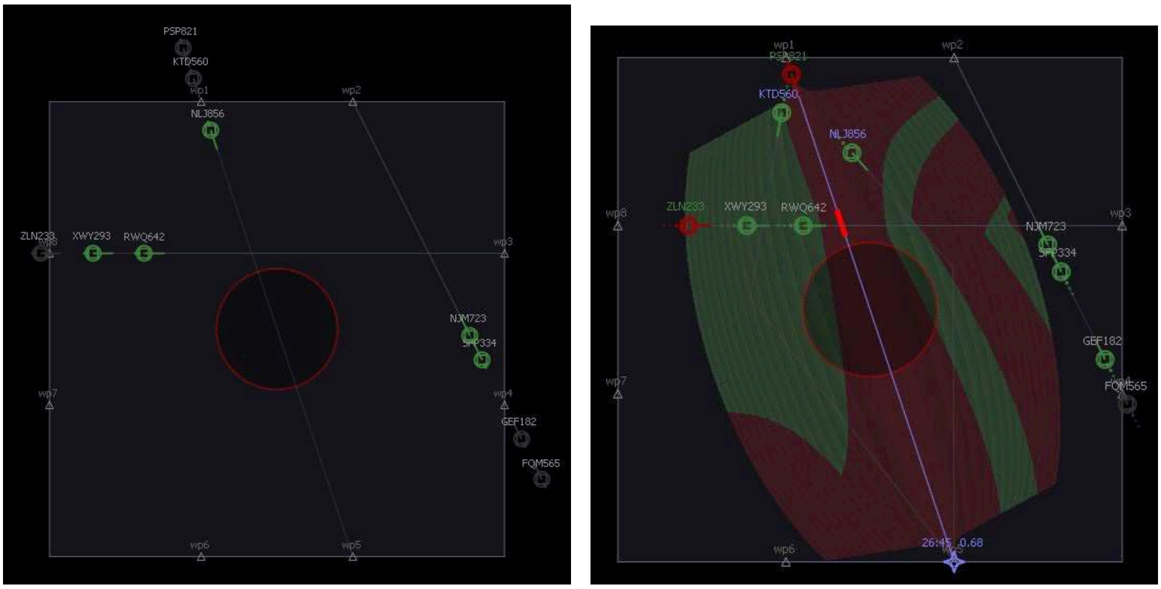


(a) Conflicts with multiple aircraft.



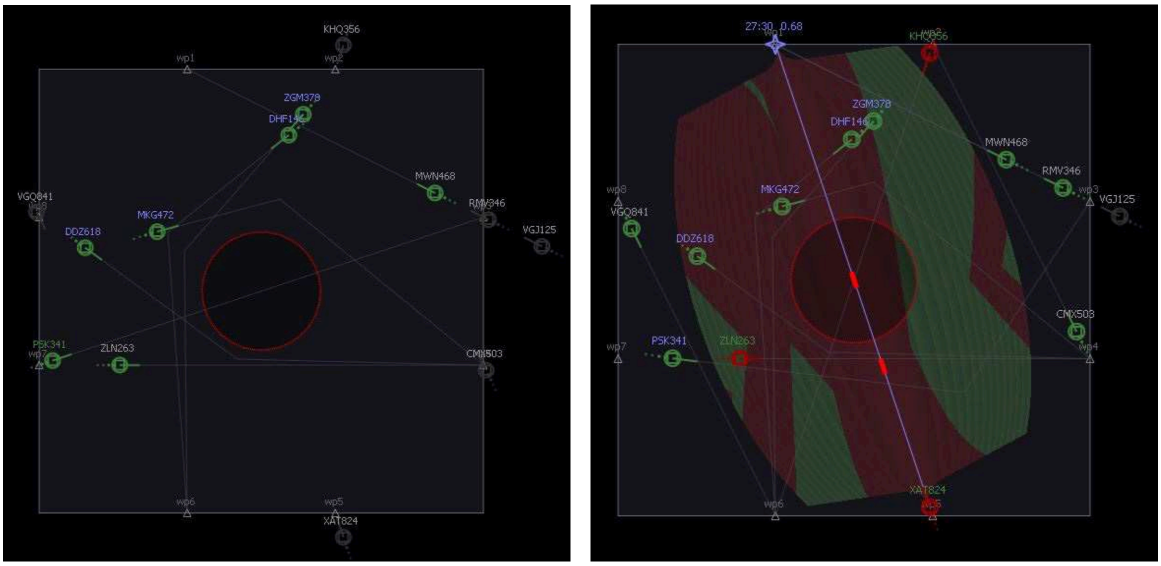
(b) Conflict resolution via waypoints.

Figure F-4: Training 4 - Conflict resolution and introduction of headphones for ISA rating.



(a) Conflicts with restricted area. **(b)** Conflict resolutions via waypoints.

Figure F-5: Training 5 - Conflict resolution, ISA rating, and restricted area.



(a) Numerous conflicts with restricted area and aircraft. **(b)** Conflict resolutions via waypoints.

Figure F-6: Training 6 - Conflict resolution, increased traffic complexity, ISA rating, and restricted area.

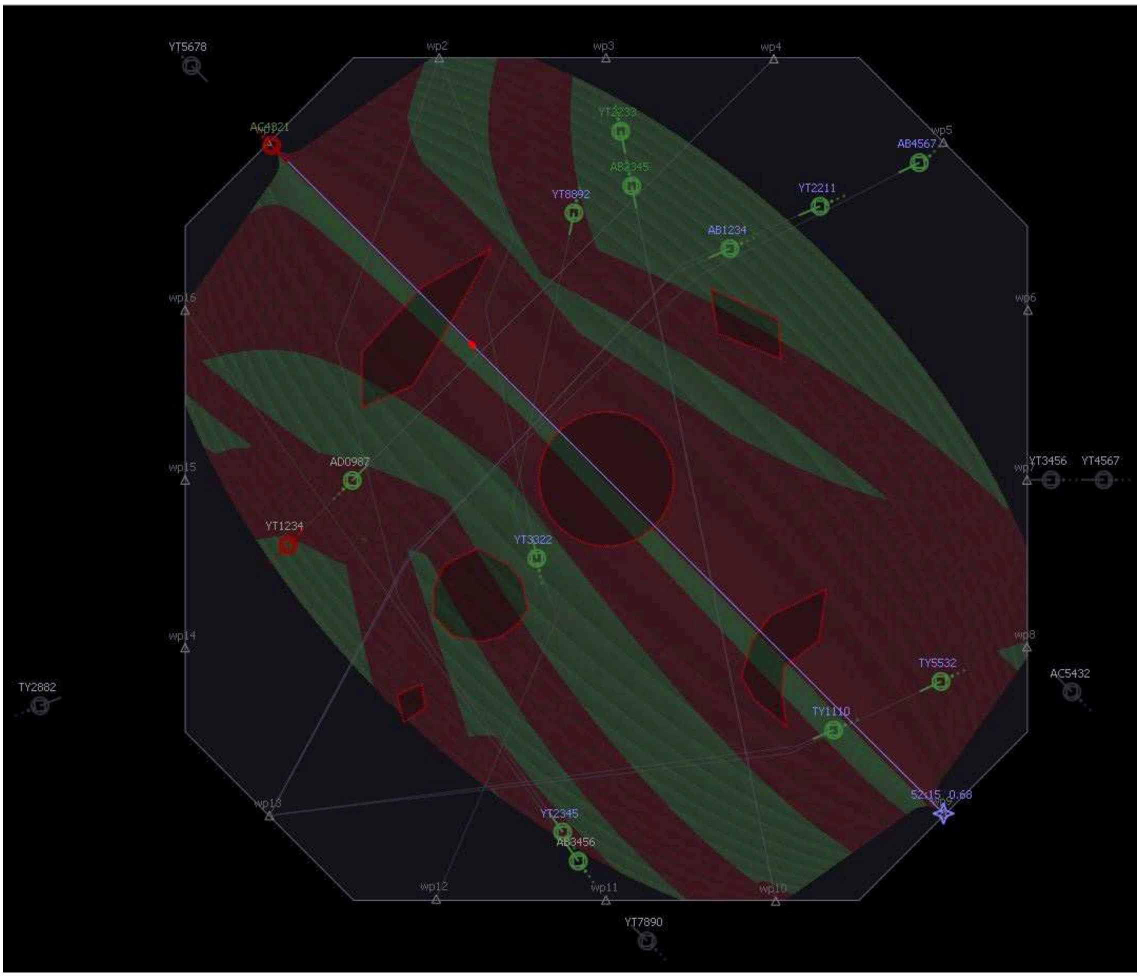


Figure F-7: Training 7 - Large sector, unstructured traffic, ISA rating, and multiple restricted areas.

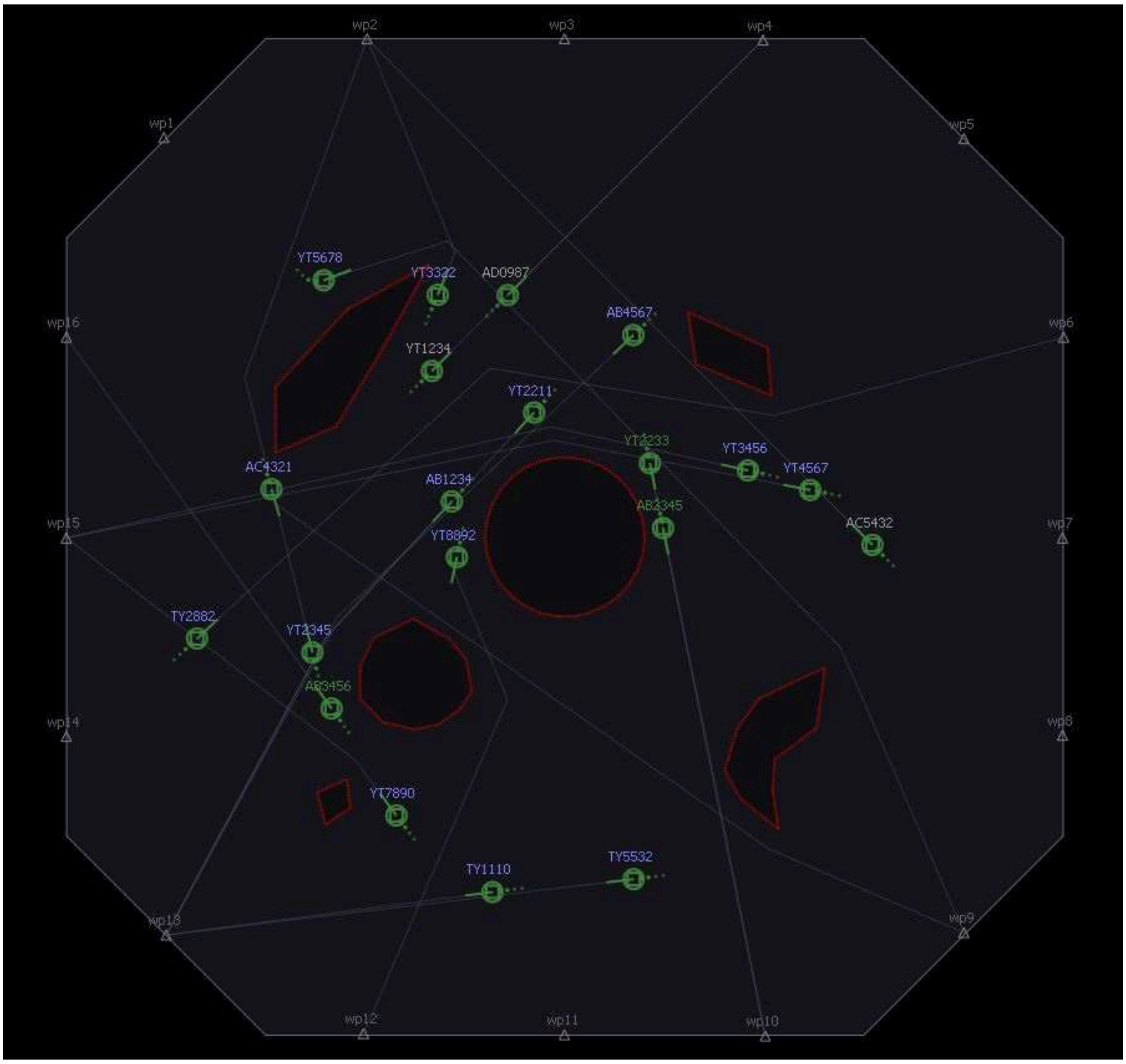


Figure F-8: Training 7 - Conflict resolutions of unstructured traffic via waypoints.

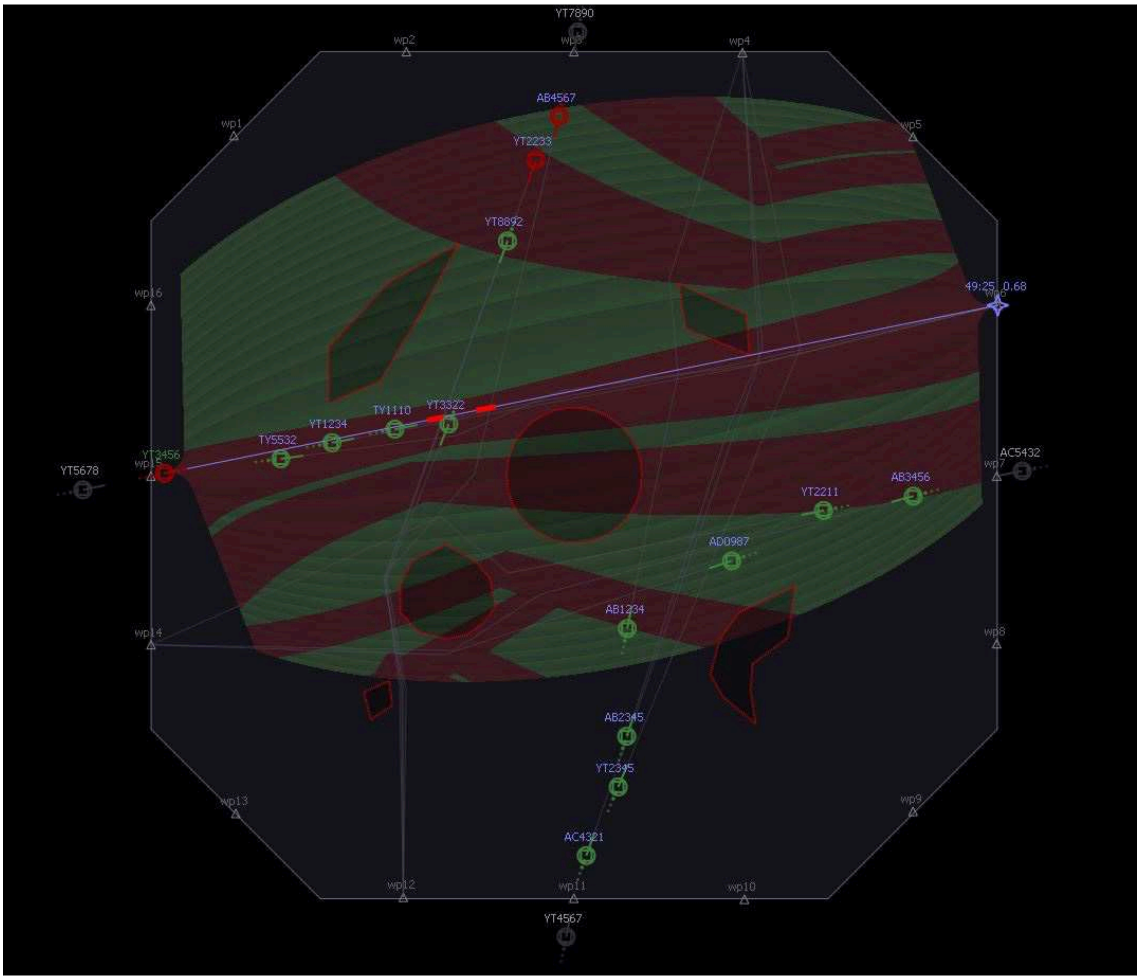


Figure F-9: Training 8 - Large sector, structured traffic, ISA rating, and multiple restricted areas.

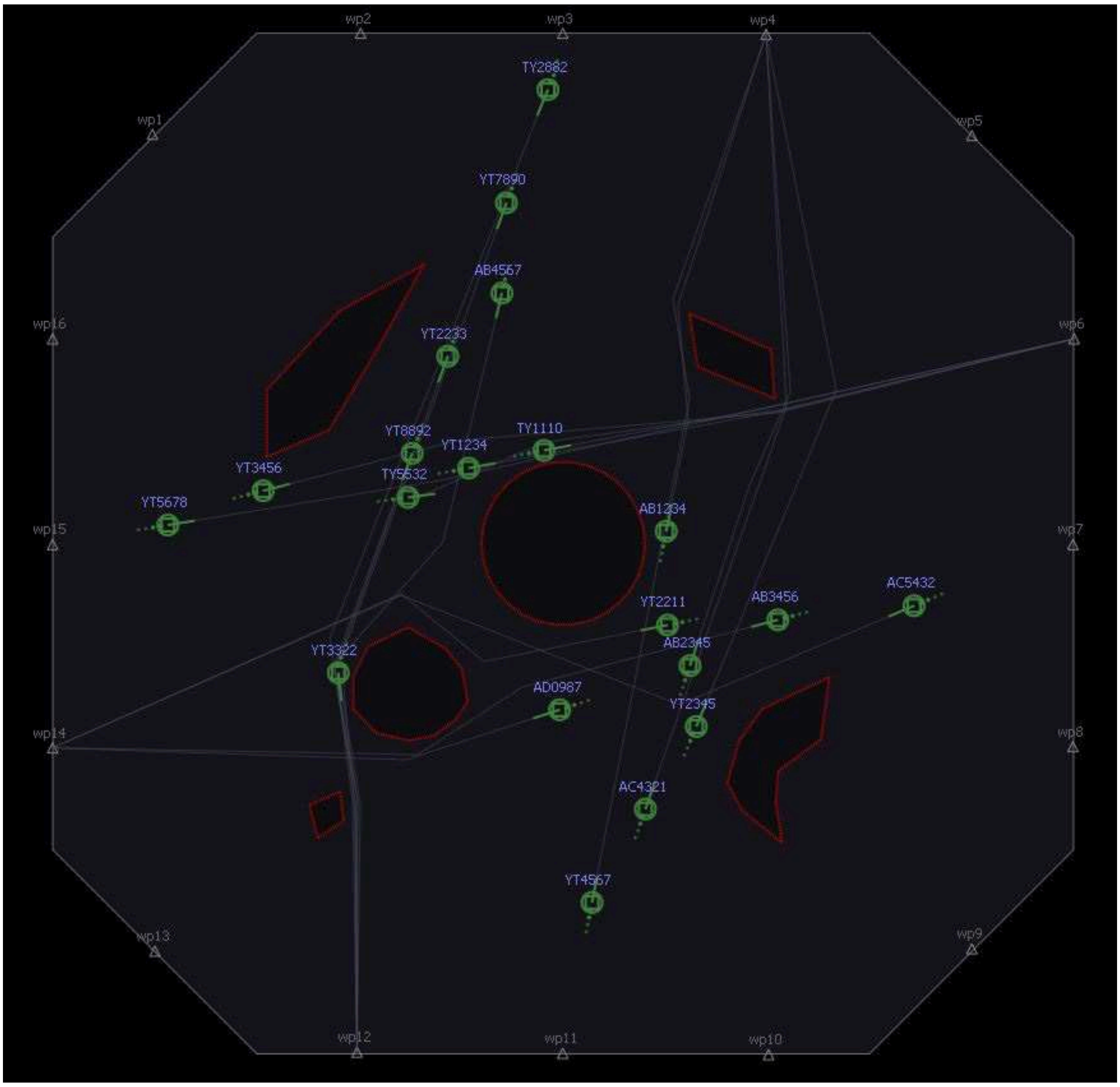


Figure F-10: Training 8 - Conflict resolutions of structured traffic via waypoints.

F-2 Experiment Scenarios

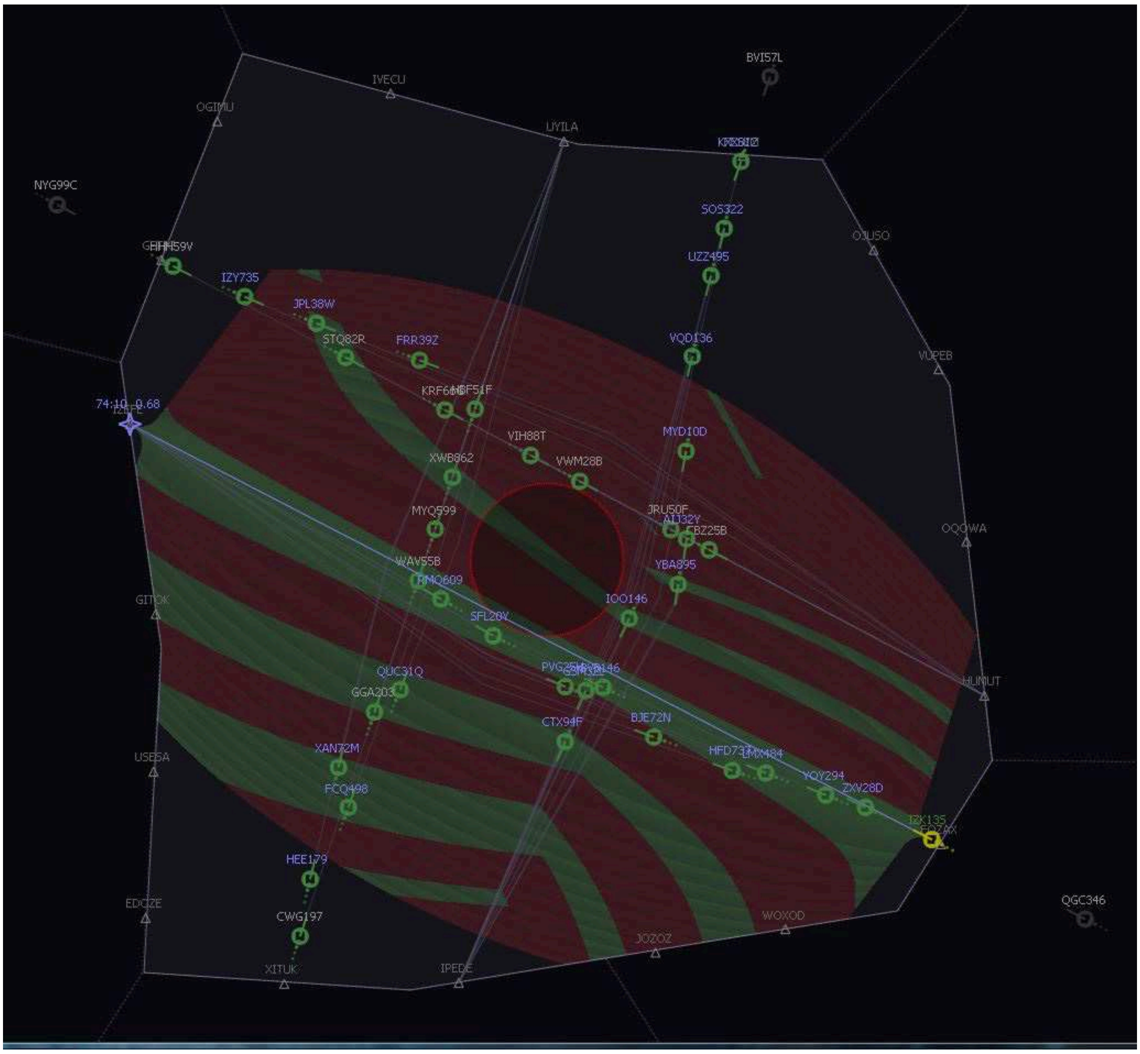


Figure F-11: S_S - Structured traffic and small perturbation.

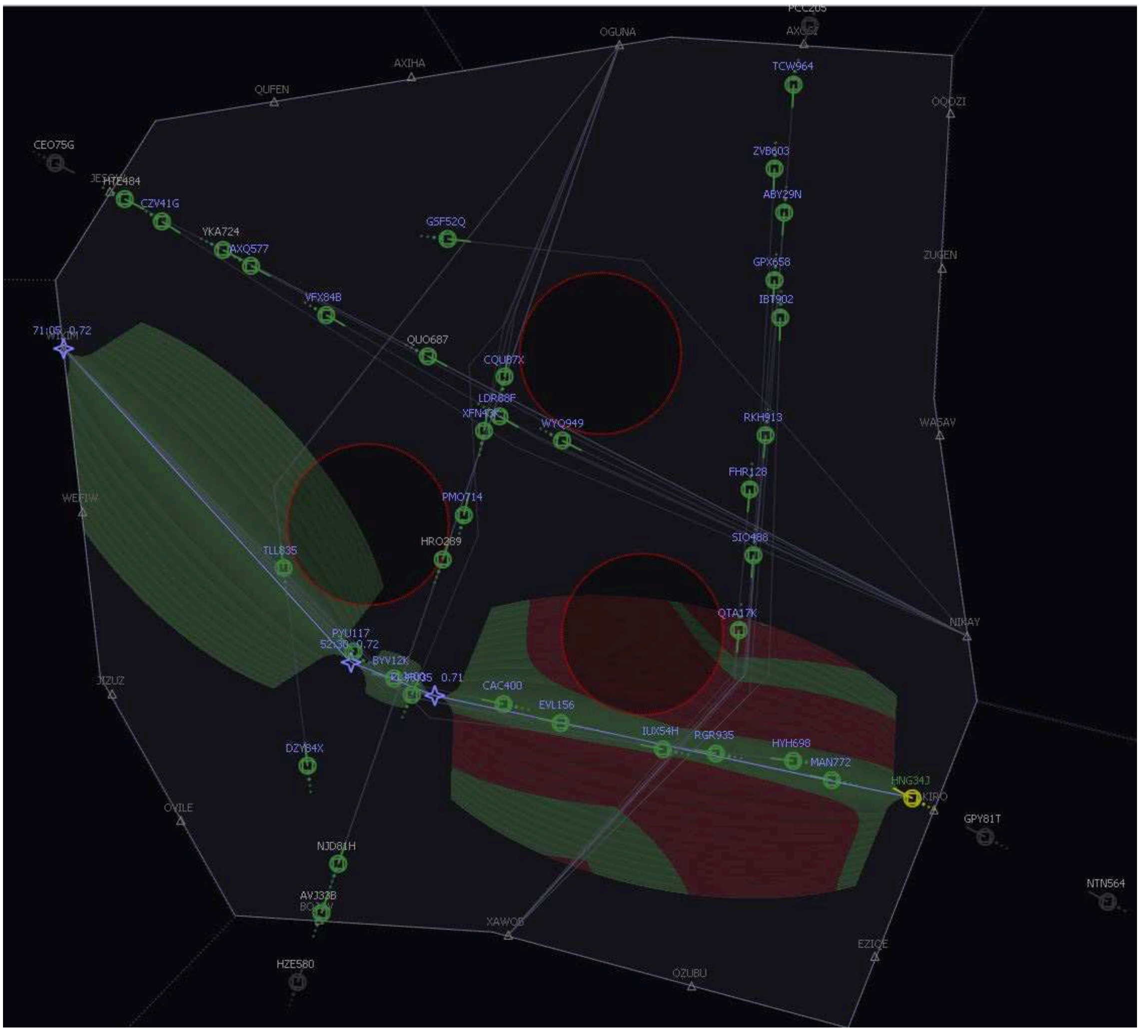


Figure F-12: S_L - Structured traffic and large perturbation.

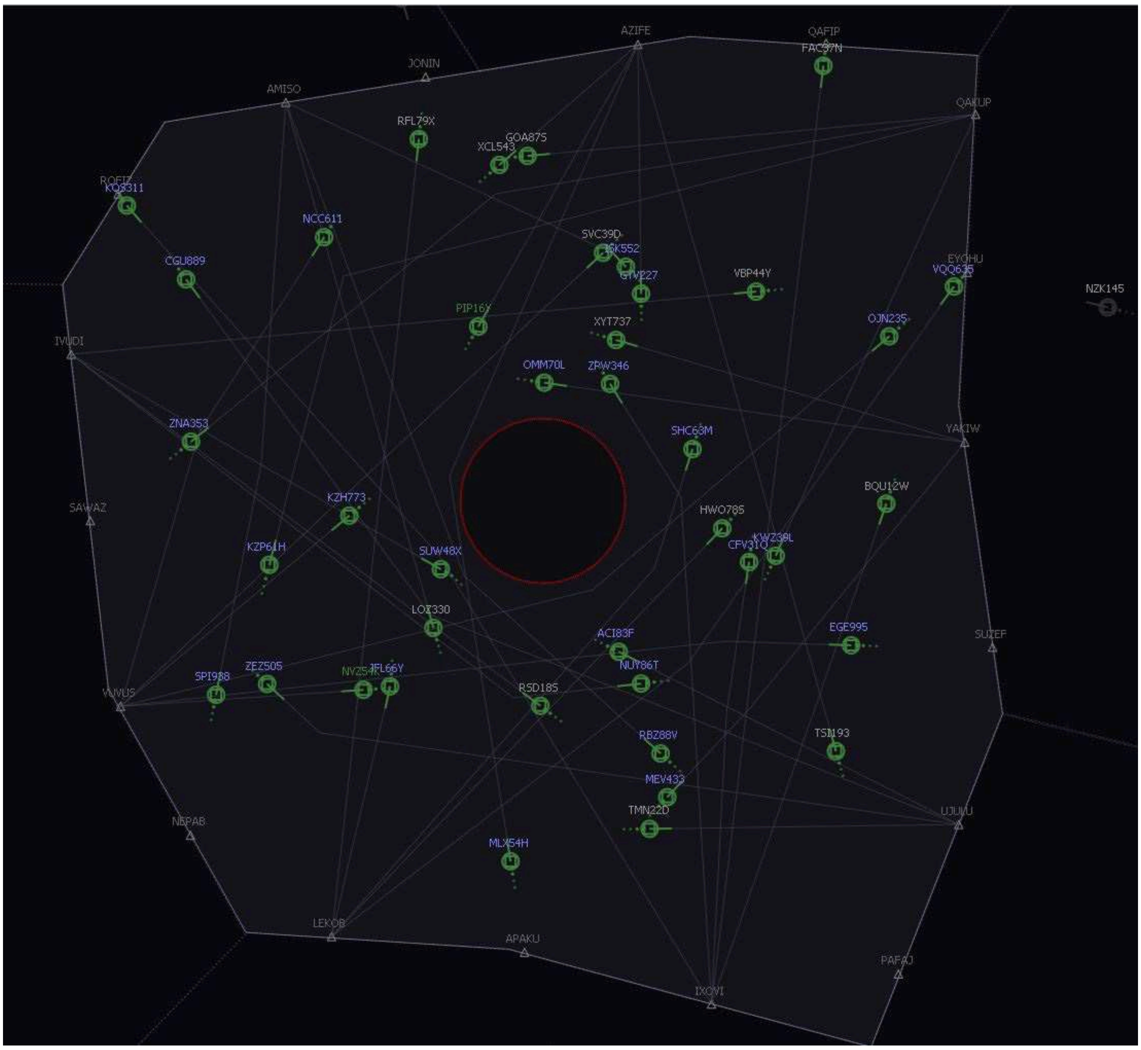


Figure F-13: U_S - Unstructured traffic and small perturbation.

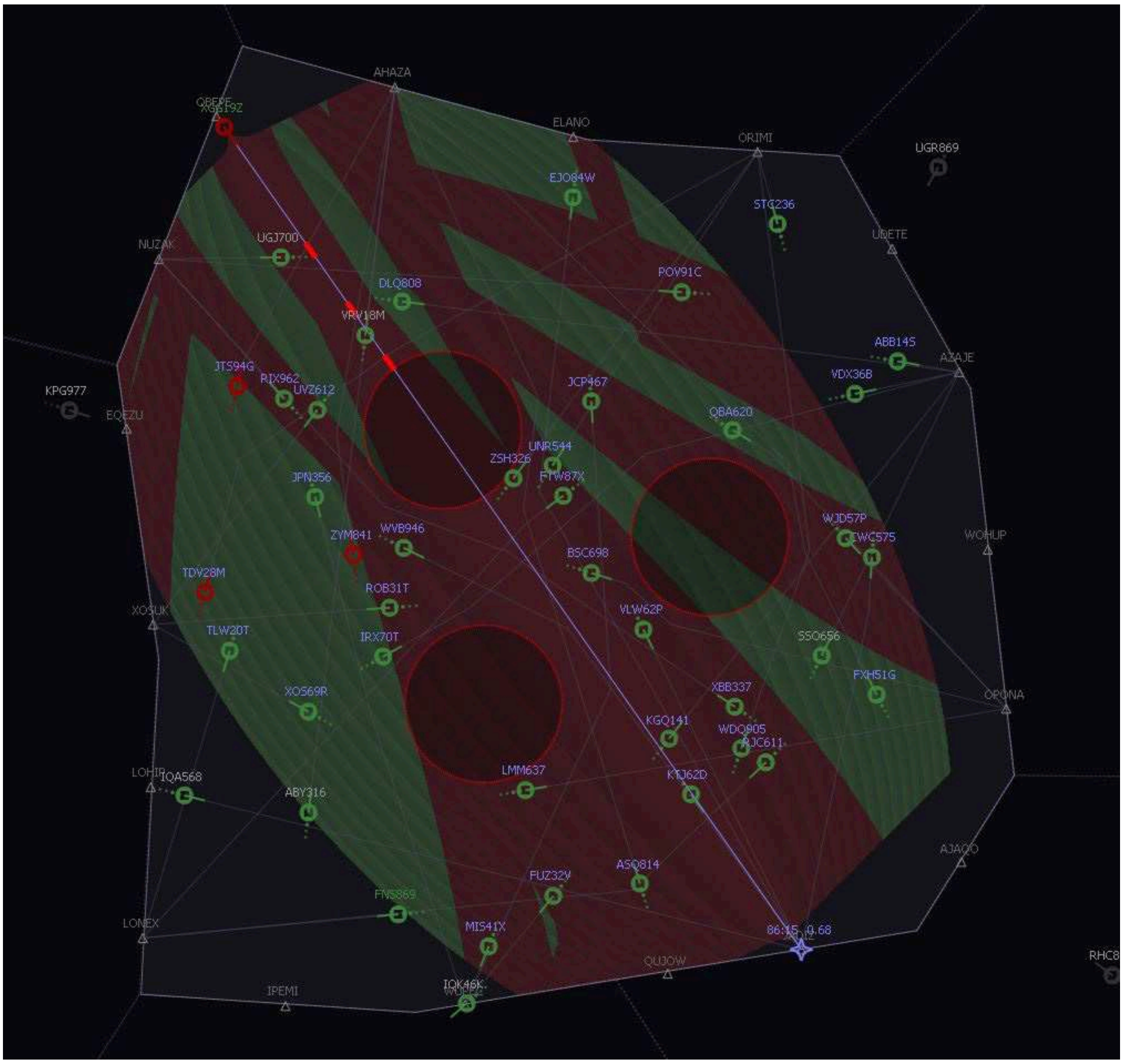


Figure F-14: U_L - Unstructured traffic and large perturbation.

Appendix G

Experiment Study Documentation

Following are the research forms and scripted prompt that were used in this study.

Information to the Participant

General information

The experiment that you are participating in today is a 4D trajectory Air Traffic Management (ATM) research tool called Travel Space being developed in the department of Control & Simulation at the TU Delft Aerospace Engineering Faculty. The aim of this research is to design a Joint Cognitive System (JCS), or human-machine ensemble, to support perturbation management in the future. This is one of many research initiatives around the world that aim to address future challenges within the ATM domain for the near future and beyond.

The experiment that you are participating in is intended to test a novel representation for supporting off nominal operations (perturbation management) in future 4D air traffic management. The aim of the experiment is thus to evaluate whether this representation can support future trajectory-based perturbation management effectively in a large airspace en-route ATM setting.

Your participation

Your participation is completely voluntary and you have the right to withdraw from the study at any moment without explanation. The recorded data are made anonymous, and are to be used solely for academic and project-related purposes.

The overall experiment

During the run of the experiment you will be asked to safely manage en-route traffic with the help of the Travel Space tool. There are four scenarios with varying traffic and restrictions placed upon the sector. Your task is to actively control the traffic by manipulating the (4D) routes of aircraft using the Travel Space representation with the help of the Travel Space tool. Please note that the system you are using presents one possible way of how aircraft might be handled in the future. This means that you might conduct tasks in a way that differs from how an en-route air traffic controller works today.

Timeline for the experiment

The overall timeline for the experiment is depicted in the table below.

Activity	Estimated duration
Introduction to the experiment	5 min
Training session	45 - 60 min
Break	10 min
Experiment run 1	15 min
Experiment run 2	15 min
Break	10 min
Experiment run 3	15 min

Experiment run 4	15 min
Debriefing	5 - 10 min

BRIEF

Training

Before the experimental runs will start, you will spend approximately 30 minutes training how to operate the system. Please make sure that you ask all questions that you have in relation with the system's functionality during the training so that you feel familiarized with the workings of the system, and feel well prepared for the experimental runs.

Scenario time & questionnaires

The scenarios during training and the experiment are in so-called scenario time, which is representing one to four times the speed of real-time. This means, for example, that each scenario in the experiment will last approximately 15 minutes, which represents about 60 minutes in scenario time.

Instantaneous Self Rating of Workload (ISA)

Every 25 seconds an Instantaneous Self Rating (ISA) Scale will pop up on the left-hand side of the screen. This scale is used to obtain your rating of the Workload experienced at that point in time of the scenario. The scale will be accompanied by an audio signal to indicate that a rating should be submitted.

Airspace & traffic

The active en-route sectors in the experiment are artificial sectors and constructed especially for this experiment. All aircraft movements are restricted to the horizontal plane (e.g., same flight level). Therefore, separating aircraft vertically will *not* be possible. All aircraft resemble a generic type of medium-sized commercial airliner and have *equal* performance (e.g., same speed range).

During the training and the experiment you are free to manage the traffic, manipulate the routes and speeds of the aircraft in whichever way you prefer.

Debrief

During the debrief session you will be free to comment on your experience of using the system.

Subject ID: ___________

Travel Space Evaluation Background Questionnaire

This questionnaire has the purpose to collect background information about the participants of the Travel Space evaluation experiment.

Age:

Gender: M / F

What is the highest degree or level of education you have completed?
If currently enrolled, mark the previous highest degree received.

- o High school degree
- o Bachelor's degree (e.g. BA, BSc)
- o Master's degree (e.g. MA, MSc, MBA)
- o Doctorate degree (PhD)
- o Other, please specify:

How regularly do you play computer, video games, or smartphone games that are related to air traffic control (ATC)?

- o Never
- o Less frequent than once a month
- o Monthly
- o Weekly
- o Daily

Please indicate which ATC game(s) and approximately how frequently you play.
(ie., hours/week, hours/month, etc.)

Game: ______/_______
Game: ______/_______
Game: ______/_______

Have you participated in any experimental study concerning Air Traffic Management prior to the Travel Space evaluation?

- o No
- o Yes, please indicate which studies:

What is your relation to the Air Traffic Control domain?

- o Actively working as an air traffic controller
- o Student within an ATC-related education program
- o Student at Aerospace Engineering
- o Student from another faculty; please specify________________________________

How would you describe your knowledge about Air Traffic Control operations?

- o Poor
- o Fair
- o Good
- o Excellent

Subject ID: _______________

Participant Consent Form

The aim of the Travel Space tool is to design a Joint Cognitive System (JCS), or human-machine ensemble, to support perturbation management in future Air Traffic Control. The aim of this experiment is thus to evaluate whether this current iteration of the Travel Space tool can support future trajectory-based perturbation management effectively in a large airspace en-route ATM setting.

During the experimental runs we will record various data. You will also be requested to answer a number of questions and questionnaires before, during and after the experiment. Your participation in this experiment is completely voluntary and you have the right to withdraw from the study at any time without having to give any explanation. In that case all data connected to you as an individual will be deleted.

Various types of data will be recorded during the experimental runs. These data, besides recordings of the traffic image will also include subjective ratings of workload, situation awareness, and controller acceptance, as well as a performance score. These will only be used for project-related documentation. Recorded data will be separated from your identity; at no time, neither now, nor in the future, will any information you provide be published that allows you as an individual to be identified. We certify to treat collected data according to good practice and follow sound ethical rules.

If you have any questions or comments concerning this study you can ask the experiment researcher.

The experiment researcher has described the purpose of the study and I know the preconditions that apply. Possible questions I had have been answered satisfactory. I am aware that behaviour related data and questionnaires will be collected and analysed. I know that I can decide to leave the experiment at any time without the need to provide any explanation.

I _____________________________ agree and participate voluntarily in this study.

Name (clear writing)

Signature: _____________________________ Date: _____________

Quick Reference Card

Aircraft Manipulation:

- select aircraft ………………………………………….. **LMB** on aircraft symbol / label
- de-select aircraft ……………………………………… **Backspace**

Trajectory Manipulation (aircraft selected):

Add Waypoint:

- add waypoint …………………………………………… **Ctrl + LMB**

Manipulate Waypoint (Waypoint Highlighted):

- move waypoint ……………………………………….. **LMB + drag**
- delete waypoint ……………………………………… **Ctrl + RMB**

Execute Trajectory (send to aircraft):

- execute ………………………………………………….. **Enter**

Aircraft Label:

Call sign	**Color status of Call sign** **Grey:** Aircraft has never been selected **Green:** Aircraft has been selected but no changes made to trajectory **Purple:** Aircraft has been selected and trajectory changes executed
Aircraft heading symbol *(solid line is the heading direction)*	**Color status of Aircraft symbol** **Green:** No Pending Aircraft Collisions *(does not account for future violations with crossing restricted airspace)* **Red:** Pending Aircraft Collisions *(does not account for future violations with crossing restricted airspace)*

Waypoint Label:

Planned fly over time *(mm:ss)*	**Required** Mach number

Training Session

It is important to note that during this training session I can answer any questions you have about the tool regarding the functionality. But once we begin the actual experiment I cannot respond to any questions you have.

Airspace and traffic

The active en-route sectors in the experiment are artificial sectors and constructed especially for this experiment. All aircraft movements are restricted to the horizontal plane (e.g., same flight level). Therefore, separating aircraft vertically will *not* be possible. All aircraft resemble a generic type of medium-sized commercial airliner and have *equal* performance (e.g., same speed range).

Training Scenario 1: System functionality and representations

1. First note that the aircraft call sign is grey, *(*physically point on the screen)* this indicates it has never been selected before. Please select the aircraft by left mouse clicking on the aircraft symbol or its call sign.
2. The aircraft symbol will turn yellow indicating that it is currently selected. The color of the call sign will change from grey (indicating that the aircraft has never been selected) into green (aircraft has been selected).
3. The representation of the aircraft is shown by the call sign and the aircraft heading symbol where the solid line is the front of the aircraft and the dots are the trailing path behind it (also see quick reference card).
4. When an aircraft is selected, its planned route is shown together with the accompanying waypoints. *(*Physically point on the screen)* Additional information is displayed in a label above each waypoint (also see quick reference card):

Planned fly over time *(mm:ss)*	Required Mach number

5. The green area displayed in front of the aircraft represents its *travel space*. The travel space visualizes *the area in which a waypoint can be added to the route segment*, not resulting in a delay to the sector exit point.
    a. De-select the aircraft by pressing the “Backspace” key or by a left click on the aircraft symbol or label. The aircraft call sign is green, indicating that the aircraft has been selected before.

Training Scenario 2: Use of waypoints

1. Please select the aircraft.
2. The red outlined area within the sector represents a restricted no fly zone, which should be avoided by all traffic.

   Note that in this experiment all aircraft will initially fly the shortest route from their sector entry to sector exit point. In case there is a restricted no fly zone along this route the aircraft will need to be actively re-routed around it. If traffic however does cross through the restricted no fly zone, a performance penalty will be given which will affect the overall performance score of the run. If an aircraft's route (grey lines) crosses a restricted no fly zone the aircraft symbol will not change to red so it is important to notice any grey route lines if they intersect a restricted no fly zone.

3. Your performance score is indicated in the lower right corner of the screen. The performance score will decrease if:
    a. An aircraft is inside a restricted no fly zone, and/or
    b. There is a conflicting pair of aircraft, which will have a loss of separation within four scenario minutes (one real time minute). This penalty increases the sooner the loss of separation will occur.
4. Please select a spot within the travel space (green area) and hold down "Ctrl."
    a. A waypoint is attached to the cursor while holding down the "Ctrl" button which can be placed anywhere within the travel space.
    b. The travel space indicates a possible solution space in which the waypoint can be added without causing a delay of the aircraft at the sector exit point.
5. Please select at spot where you would like to add a waypoint by holding down the "Ctrl" button and left click to add a waypoint to re-route the aircraft around the restricted no fly zone.
    a. Now your waypoint has been added, the planned route has been split it into two segments with *equal speed*. The speed is set such that the aircraft will still arrive at its sector exit point at the planned time. Furthermore, each segment now has its own travel space indicating where another additional waypoint can be placed. You can place multiple waypoints down at a time while still holding down the "Ctrl" button.
    b. After adding the waypoint it can still be dragged in case that you would like to reposition it. This is done by holding the left click and dragging it to the desired position. Note that once the waypoint is added its time is fixed. As a consequence, dragging it will result in a change of the required speeds for the adjacent segments. Please drag the waypoint slightly to see this effect.

       Note that until now, all modifications to the aircraft route are not yet sent to the aircraft itself. In this "probing phase" all modifications can still be un-done by deselecting the aircraft.
6. Push the "Enter" key to execute.
    a. Through executing, all changes will be made permanent and the aircraft will start to execute the new plan.
    b. Note that the color of the aircraft call sign has turned purple. This indicates that it is flying along an updated plan from the controller.
7. Please add another waypoint.
    a. Additionally, waypoints can be deleted by while holding down "Ctrl" and right clicking on that waypoint.
8. Please delete the waypoint you just added.

Training Scenario 3: Conflict resolution

All aircraft that are currently recognized as being involved in a conflict will be displayed red instead of green.

1. Select an aircraft in conflict.
    a. The red emphasis or red highlighted areas *(*physically point at these)* on its route indicates the presence and location of a future loss of separation between aircraft. In this experiment conflicts should foremost be resolved through the introduction of one or more waypoints.

   b. Note that in the right top corner a table displays all conflict pairs. Each table entry indicates the call signs of the conflict pair, closest point of approach (in NM) and time to loss of separation *(mm:ss)*. The table entries are sorted by time to loss of separation.
   c. Next to the green areas in the travel space as seen previously, red areas now indicate no-go positions for waypoint placement. Placing a waypoint in such a zone will result in a new conflict -or conflicts- with other traffic.
2. Introduce one or more waypoints to the route of one or more aircraft to resolve all indicated conflict(s).

Scenario 4: Conflict resolution

While all previous training scenarios have been presented in scenario time, this scenario will simulate the traffic in "real time", which represents 4 x scenario time speed. Before we start the training scenario I will explain this training to you, as you will be putting on headphones later.

Instructions

1. Please resolve the conflict(s) present in the scenario.
2. Every 25 seconds (real-time), a 0-100 scale is shown on the left hand side of the display. This event is accompanied with an audible queue. With this so-called ISA-scale (Instantaneous Self-Assessment of Workload), you are asked to rate your current workload while performing the experiment task. Where 0/green stands for a low workload and 100/red for high workload.
   a. Please rate your experienced workload on the scale on the left hand side of the screen by left mouse clicking. It is very important to complete this rating scale as quickly and accurately as you can once it pops up as this data will help in assessing the tool later. Please put on the HEADPHONES NOW.

Scenario 5: Repetition

The scenario will run in real time (4x scenario time).

Instructions

1. Please resolve the conflict(s) present within the scenario. The restricted no fly zone restricts your control actions; aircraft must not enter this area.

   If, for any reason, the conflict cannot be resolved without crossing the special use cell, you will receive a performance penalty.
2. Please remember to rate your experienced workload on the ISA-scale.

Training: Scenario 6 Repetition

The scenario will run in scenario time (4x real time)

Instructions

1. Please resolve the conflict(s) present in the scenario.
2. Please remember to rate your experienced workload on the ISA-scale.

Training: Scenario 7 Repetition Large Scale Unstructured Traffic

The scenario will run in scenario time (4x real time)

Instructions

1. Please resolve the conflict(s) present in the scenario.
2. Please remember to rate your experienced workload on the ISA-scale.

Training: Scenario 8 Repetition Large Scale Structured Traffic

The scenario will run in scenario time (4x real time)

Instructions

1. Please resolve the conflict(s) present in the scenario.
2. Please remember to rate your experienced workload on the ISA-scale.

Training: Repetition if Needed

Repeat Scenario 7 and/or 8 if participant does not seem to grasp the hang of using the travel space tool yet. Do it until they feel comfortable using the tool and they are making multiple waypoints to resolve conflicts around restricted areas as well as LOS. Also until they get use to noticing the grey route lines crossing the restricted areas.

Experiment Training Overview

Aim

The aim of this training is to make sure that possible interactions have been tested by the research participants.

- Understand the information presented within the system
- Understand the representation of a single aircraft
- Understand how perturbations are represented
- Understand how the travel space can be used as a decision support for conflict resolution
- Be able to solve conflict situations due to minor or larger perturbations by re-routing one or more aircraft and/or change sector exit times

Scenario	Aim	Time	Comments
Scen 1	Introduce system and representations	2-4 min	Speed 1x scenario time
Scen 2	Introduce special use airspace & use waypoints	2-4 min	Speed 1x scenario time
Scen 3	Conflict resolution, small perturbation	2-4 min	Speed 1x scenario time
Scen 4	Repetition, ISA first time, HEADPHONES!	3-4 min	Speed 4x scenario time
Scen 5	Repetition, ISA	3-4 min	Speed 4x scenario time
Scen 6	Repetition, ISA, increased traffic complexity	5-10 min	Speed 4x scenario time
Scen 7	Repetition, UNSTRUCTURED traffic, increased airspace size and various restricted airspaces, ISA	6-10 min	Speed 4x scenario time
Scen 8	Repetition, STRUCTURED traffic, increased airspace size and various restricted airspaces, ISA	6-10 min	Speed 4x scenario time
Repetition Scen 7/8	Repeat if needed until familiar and comfortable with tool	6-20 min	Speed 4x scenario time

Appendix H

Experiment Study Results of Complexity Metric and ISA Plots

H-1 Dynamic Density vs Instantaneous Self Rating

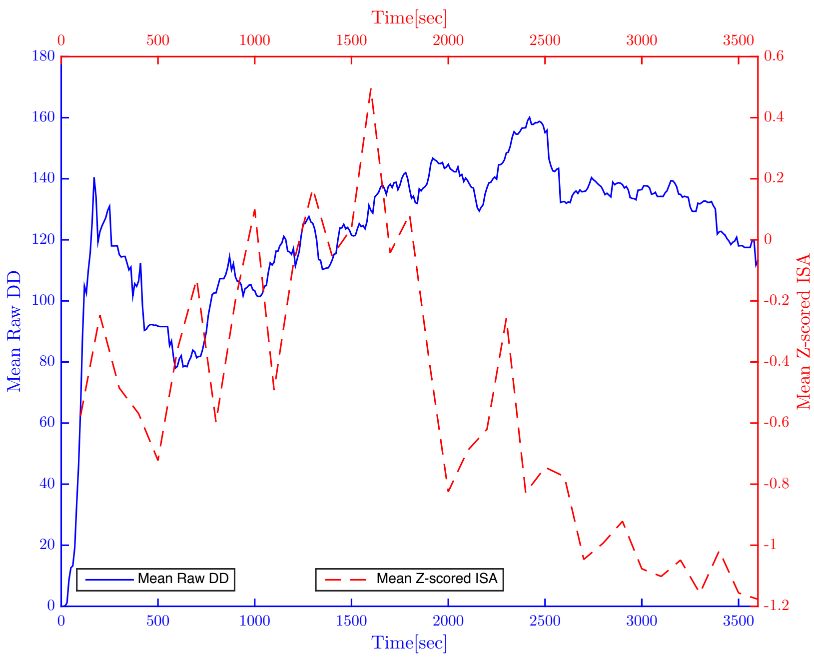


Figure H-1: Experiment Z-scored mean across all participants of DD with ISA - Structured Traffic, Small Perturbation

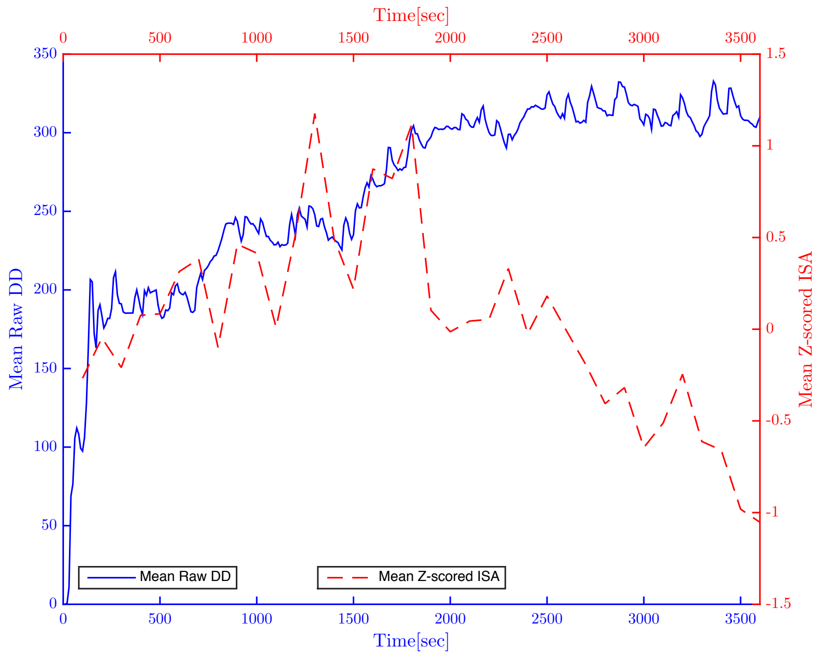


Figure H-2: Experiment Z-scored mean across all participants of DD with ISA - Structured Traffic, Large Perturbation

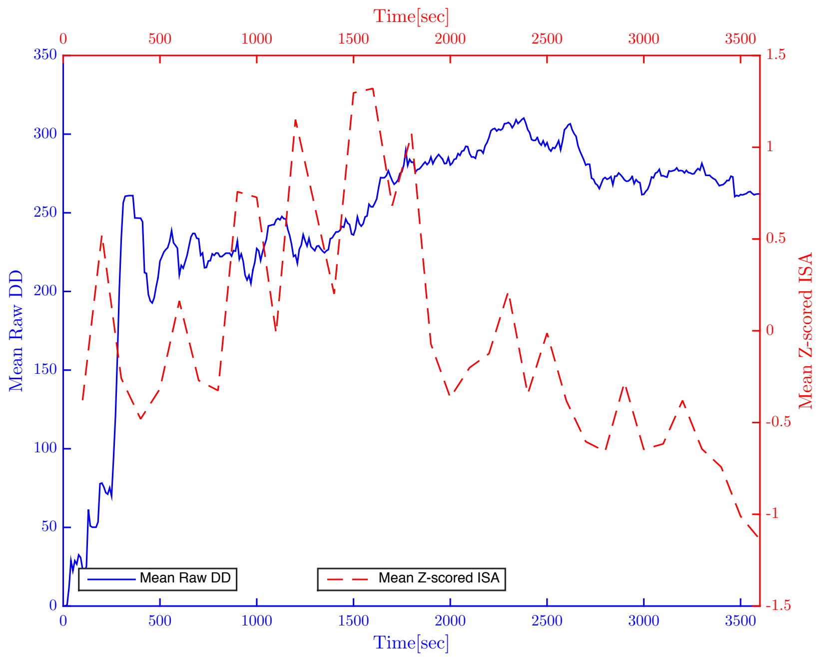


Figure H-3: Experiment Z-scored mean across all participants of DD with ISA - Unstructured Traffic, Small Perturbation

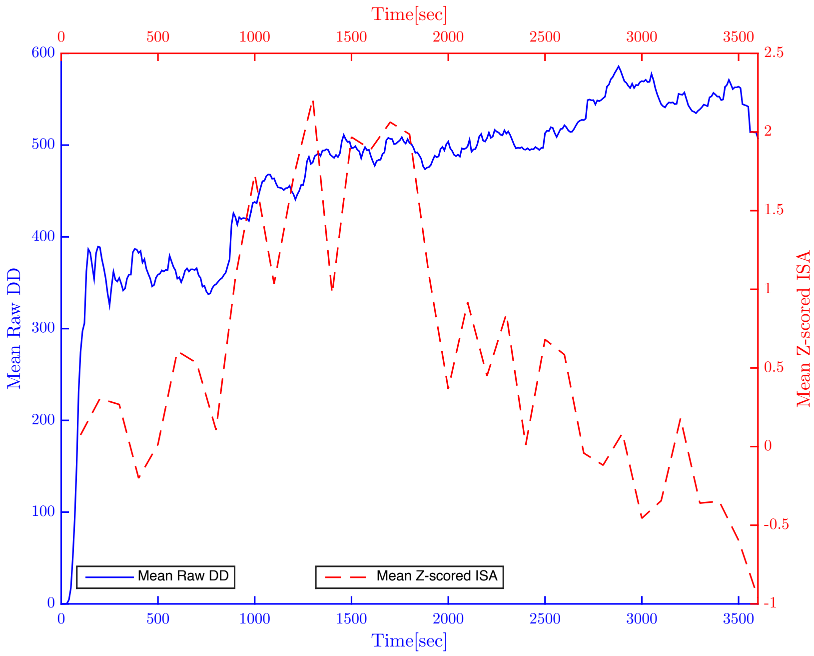


Figure H-4: Experiment Z-scored mean across all participants of DD with ISA - Unstructured Traffic, Large Perturbation

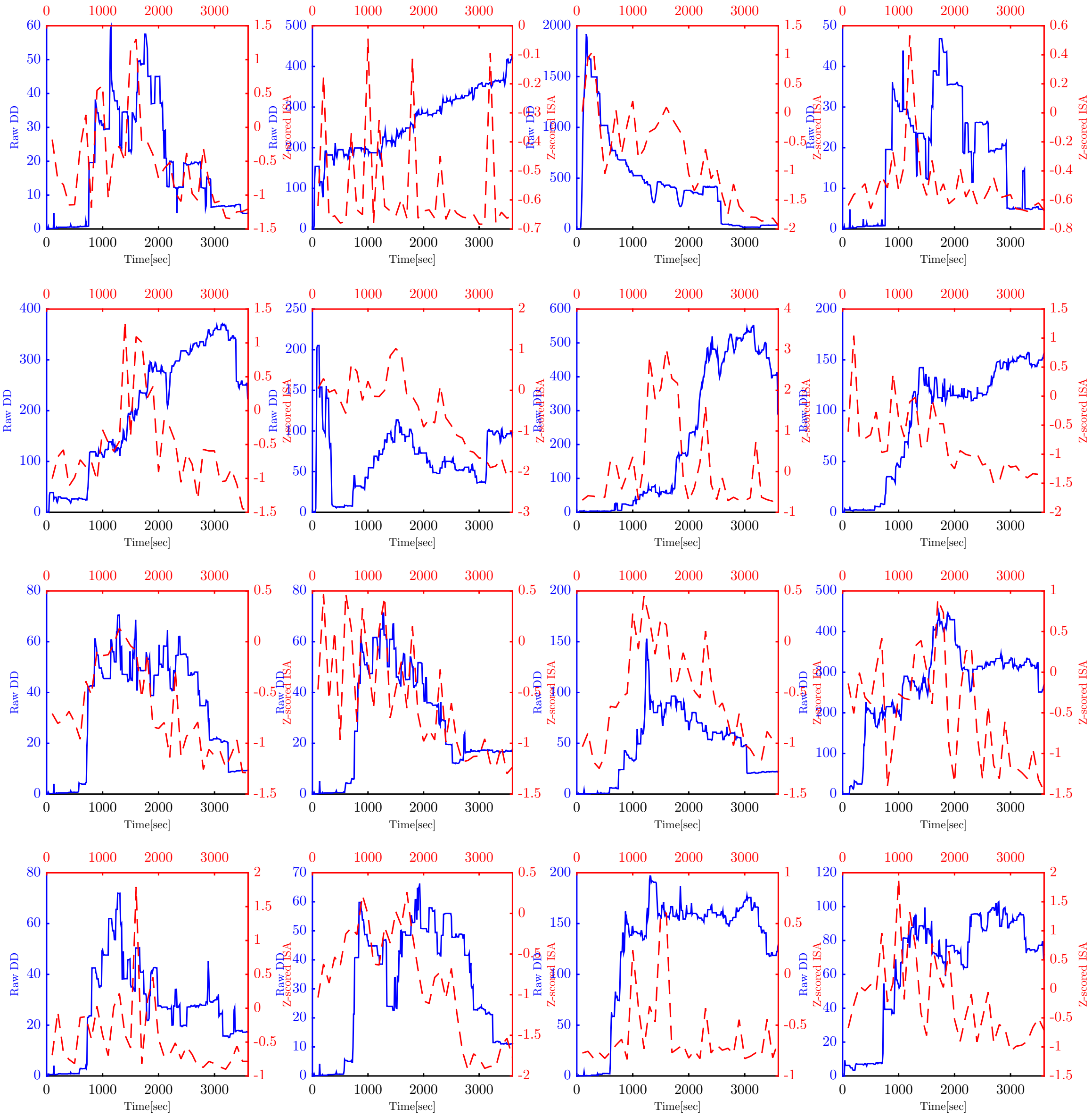


Figure H-5: All individual participant results of DD with ISA - Structured Traffic, Small Perturbation

Figure H-6: All individual participant results of DD with ISA - Structured Traffic, Large Perturbation

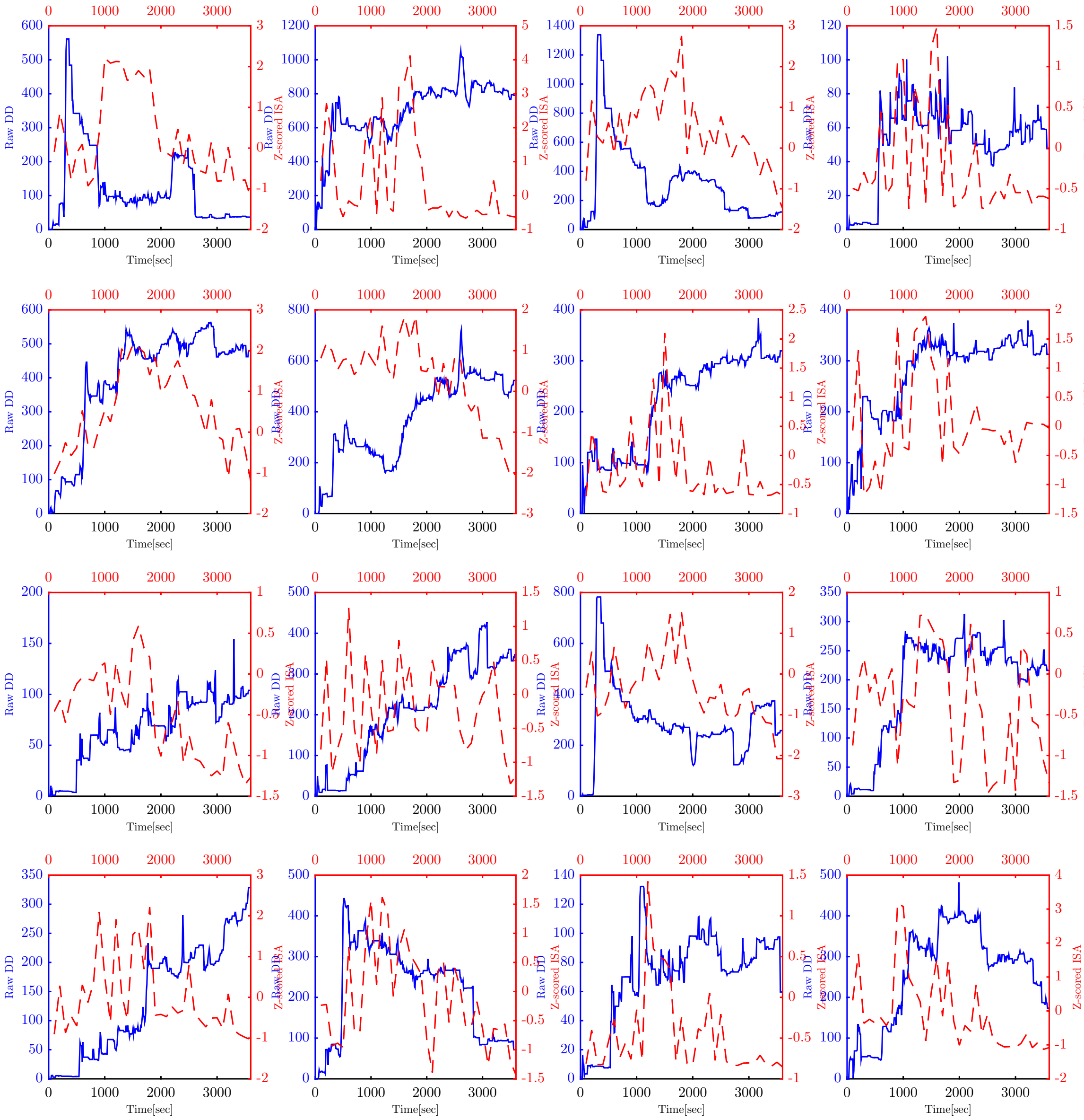


Figure H-7: All individual participant results of DD with ISA - Unstructured Traffic, Small Perturbation

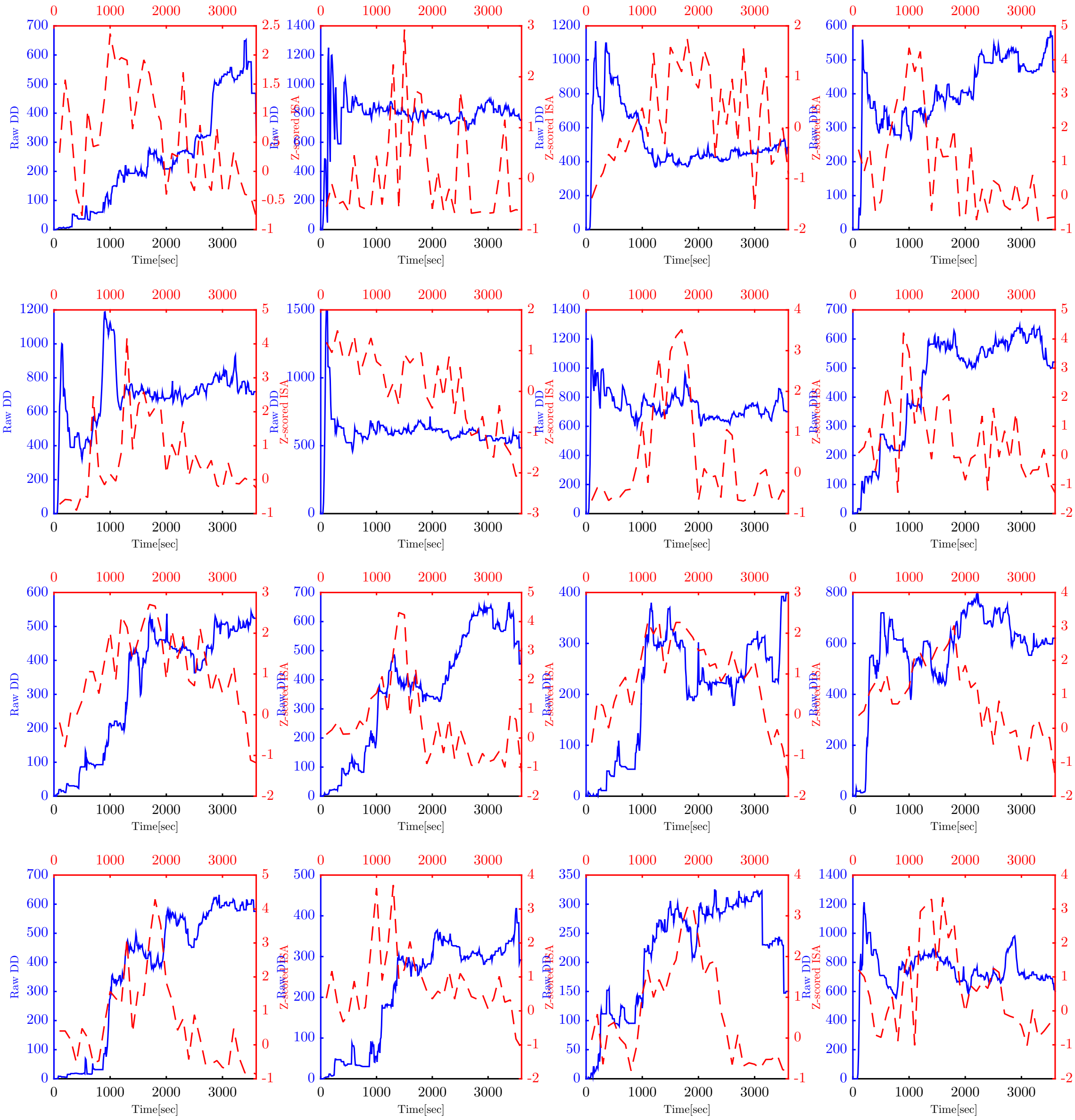


Figure H-8: All individual participant results of DD with ISA - Unstructured Traffic, Large Perturbation

H-2 Trajectory-based Complexity vs Instantaneous Self Rating

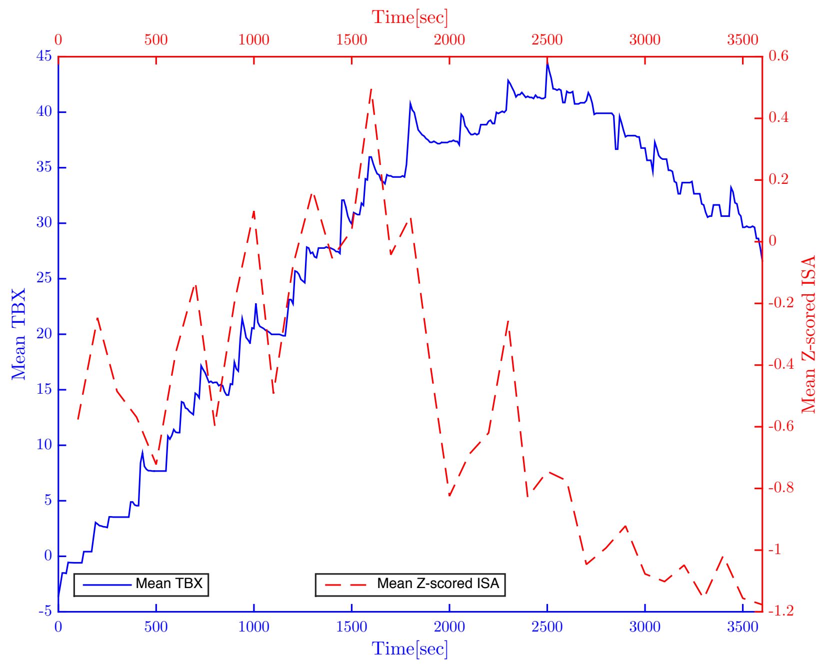


Figure H-9: Experiment Z-scored mean across all participants of TBX with ISA - Structured Traffic, Small Perturbation

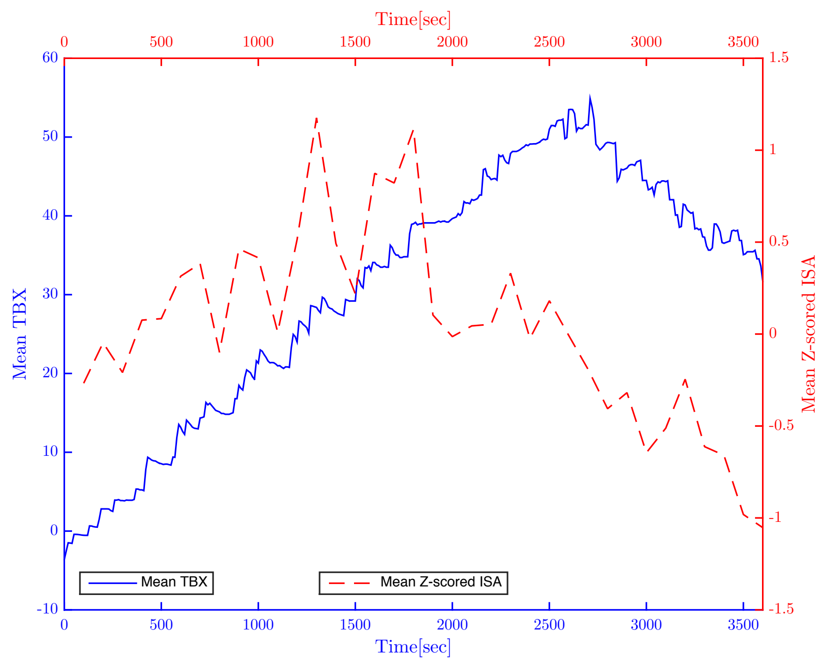


Figure H-10: Experiment Z-scored mean across all participants of TBX with ISA - Structured Traffic, Large Perturbation

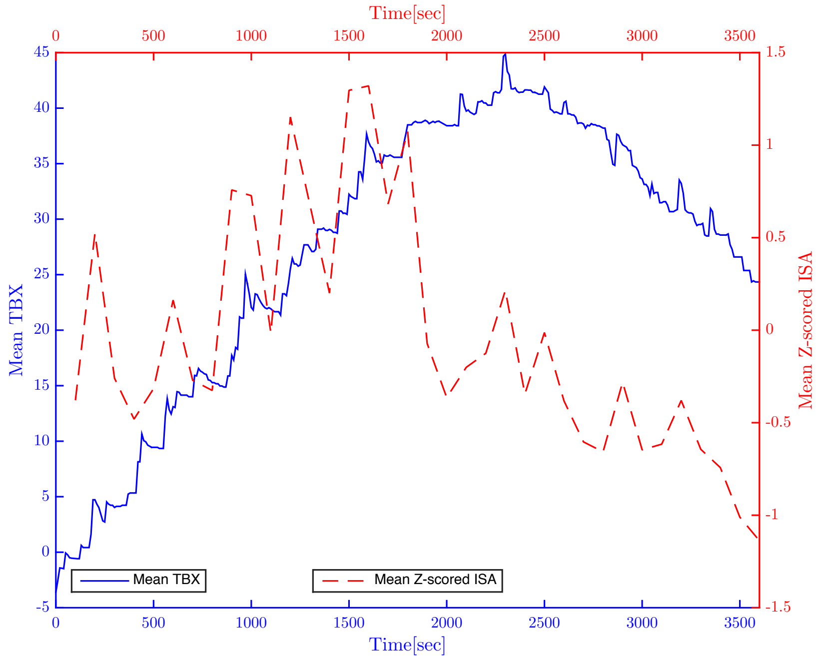


Figure H-11: Experiment Z-scored mean across all participants of TBX with ISA - Unstructured Traffic, Small Perturbation

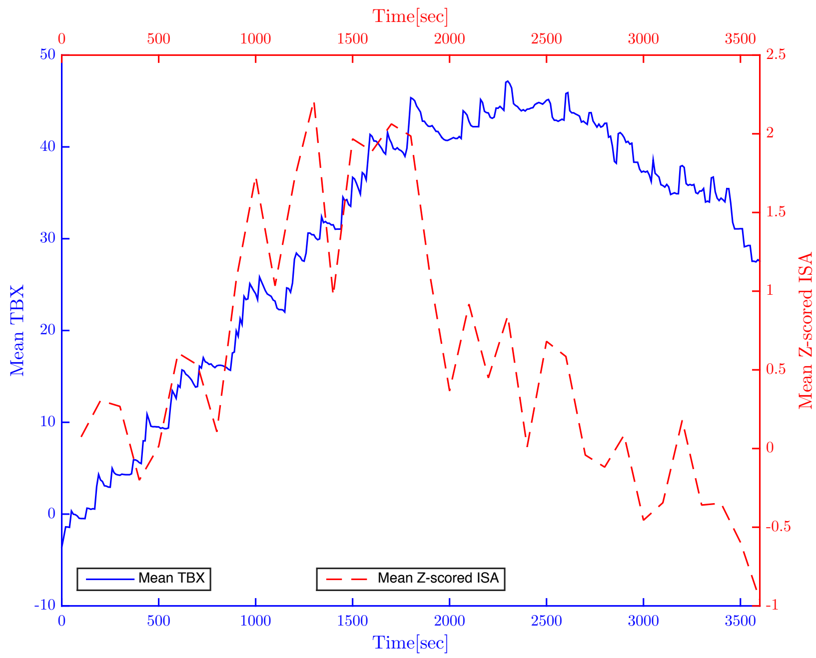


Figure H-12: Experiment Z-scored mean across all participants of TBX with ISA - Unstructured Traffic, Large Perturbation

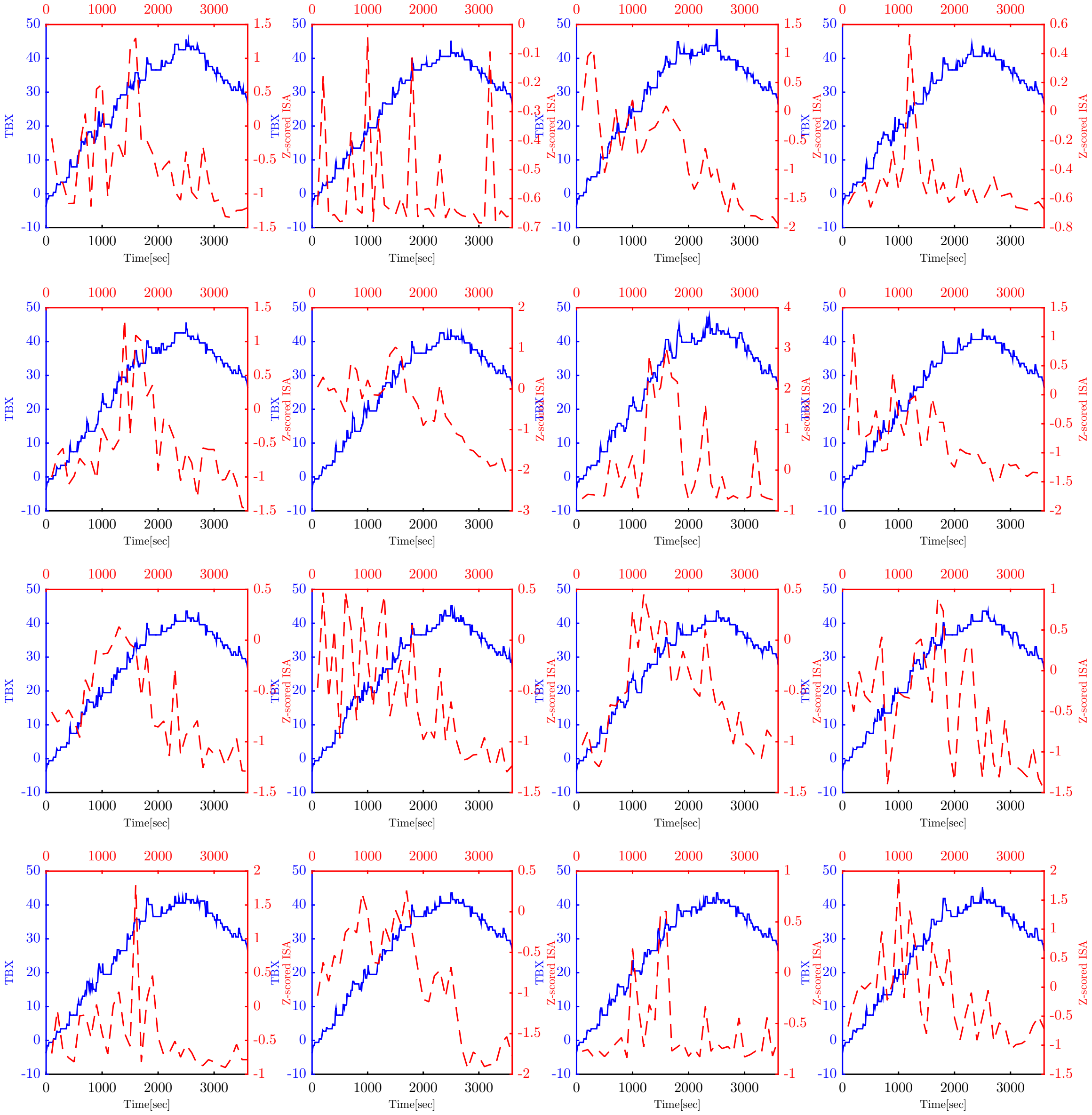


Figure H-13: All individual participant results of TBX with ISA - Structured Traffic, Small Perturbation

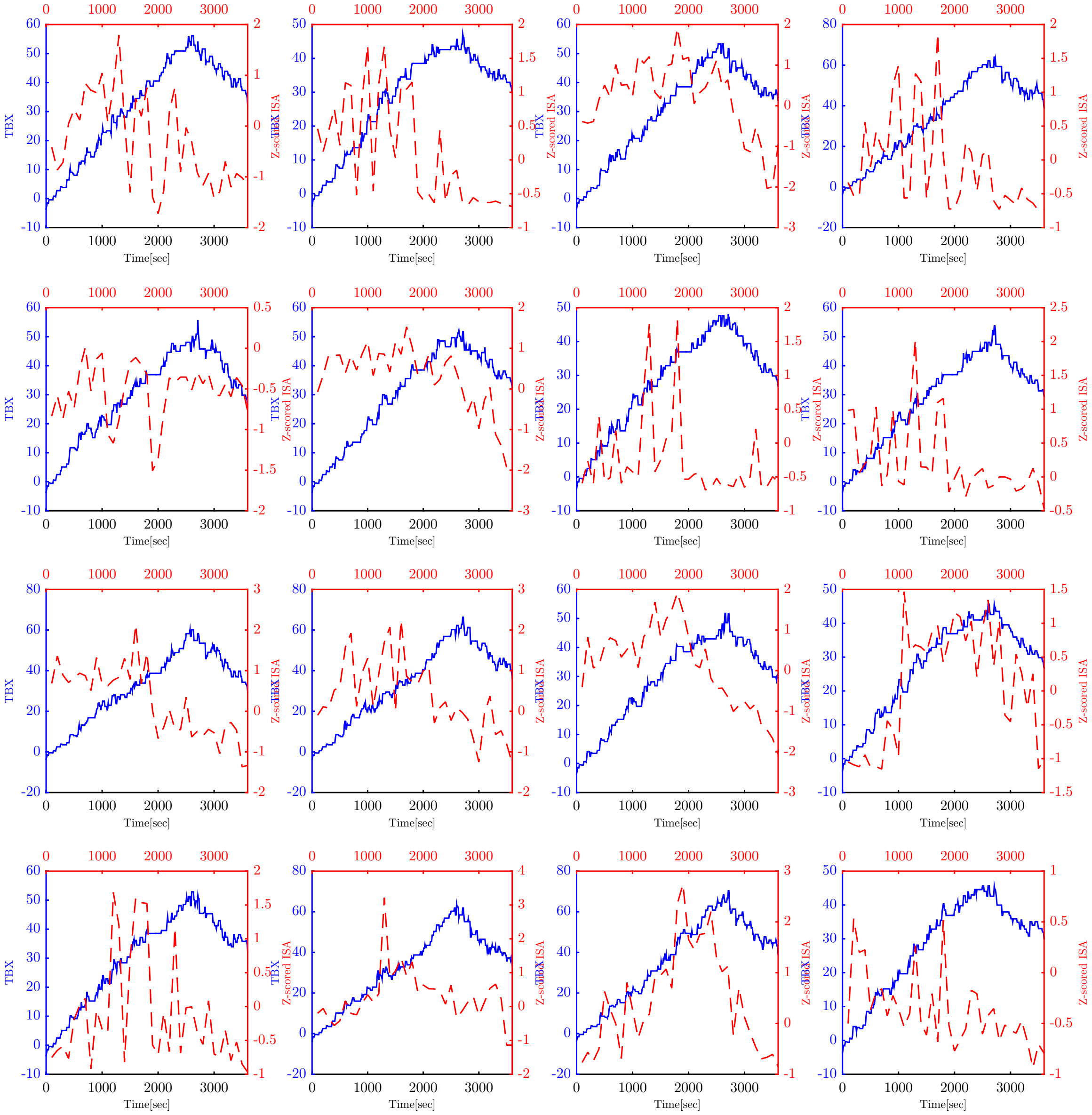


Figure H-14: All individual participant results of TBX with ISA - Structured Traffic, Large Perturbation

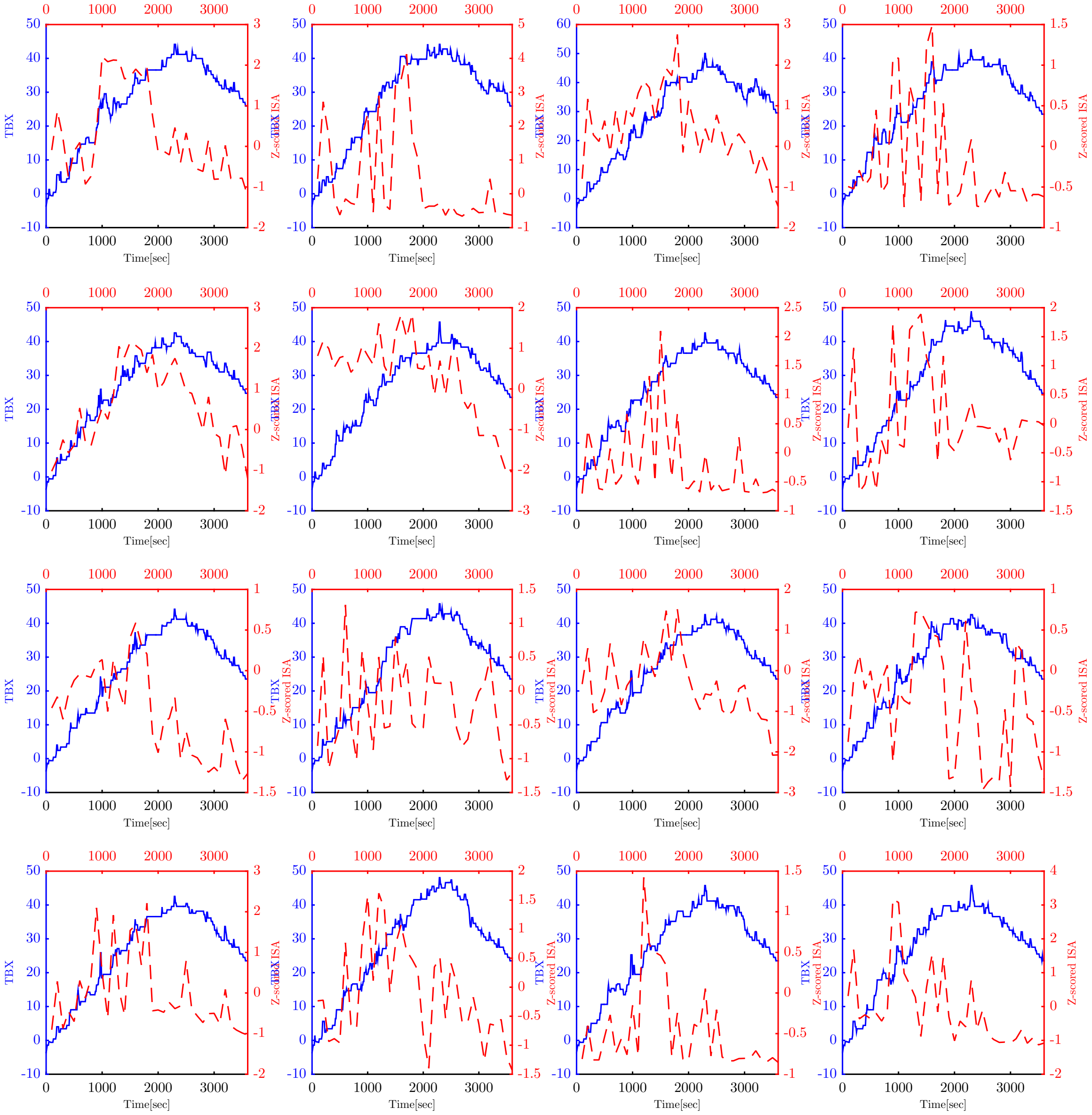


Figure H-15: All individual participant results of TBX with ISA - Unstructured Traffic, Small Perturbation

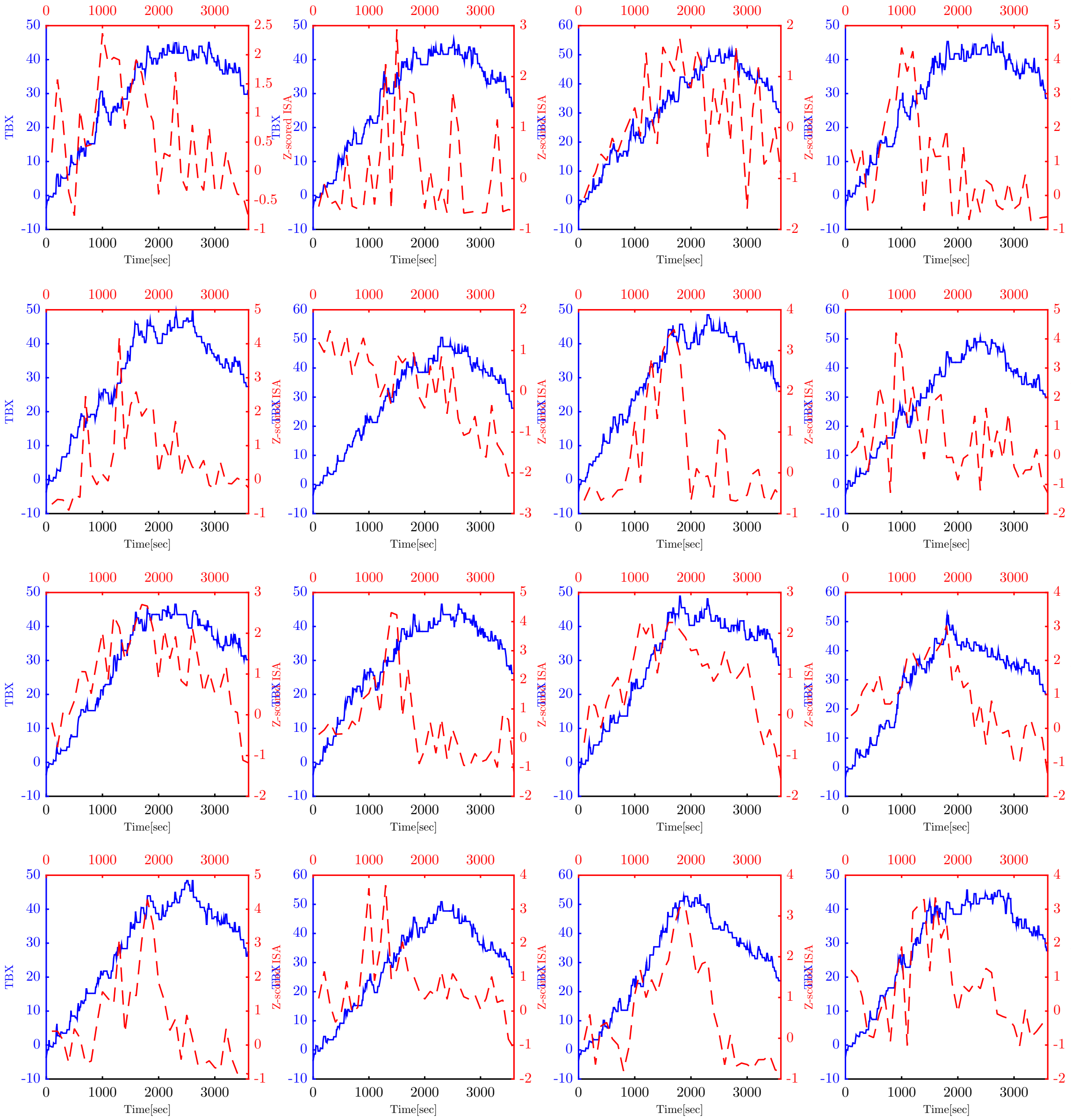


Figure H-16: All individual participant results of TBX with ISA - Unstructured Traffic, Large Perturbation

H-3 Time Shifted Trajectory-based Complexity vs Instantaneous Self Rating

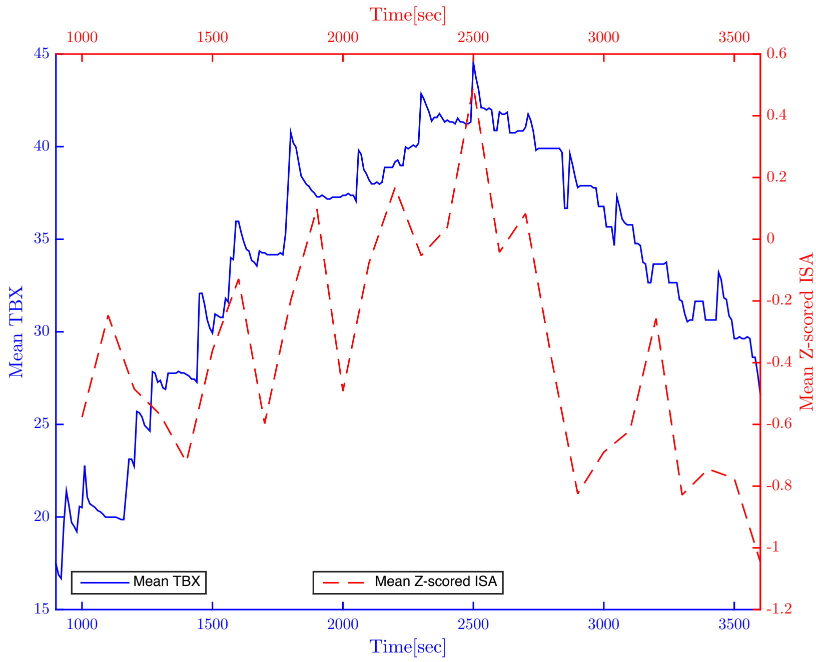


Figure H-17: Experiment time shifted 900 seconds, Z-scored mean across all participants of TBX with ISA - Structured Traffic, Small Perturbation

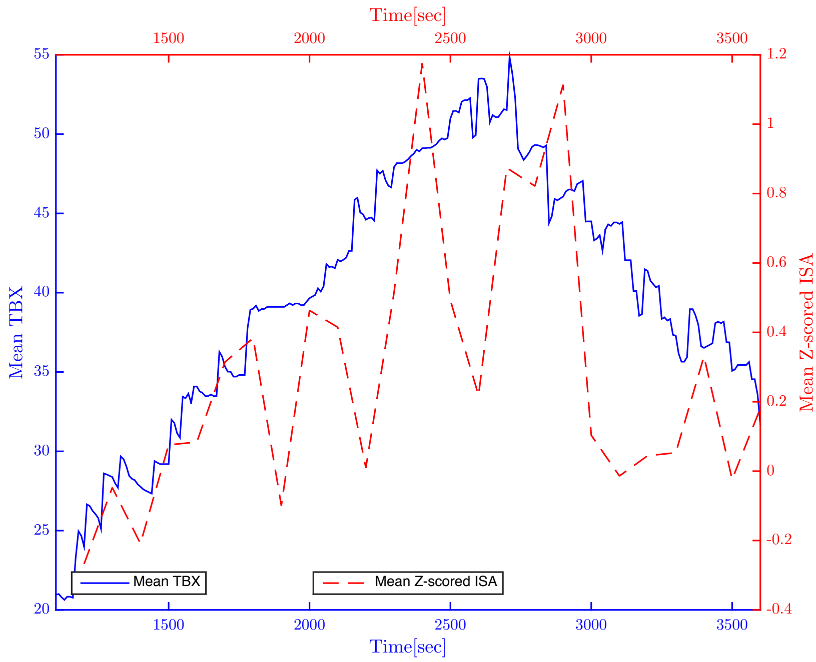


Figure H-18: Experiment time shifted 1100 seconds, Z-scored mean across all participants of TBX with ISA - Structured Traffic, Large Perturbation

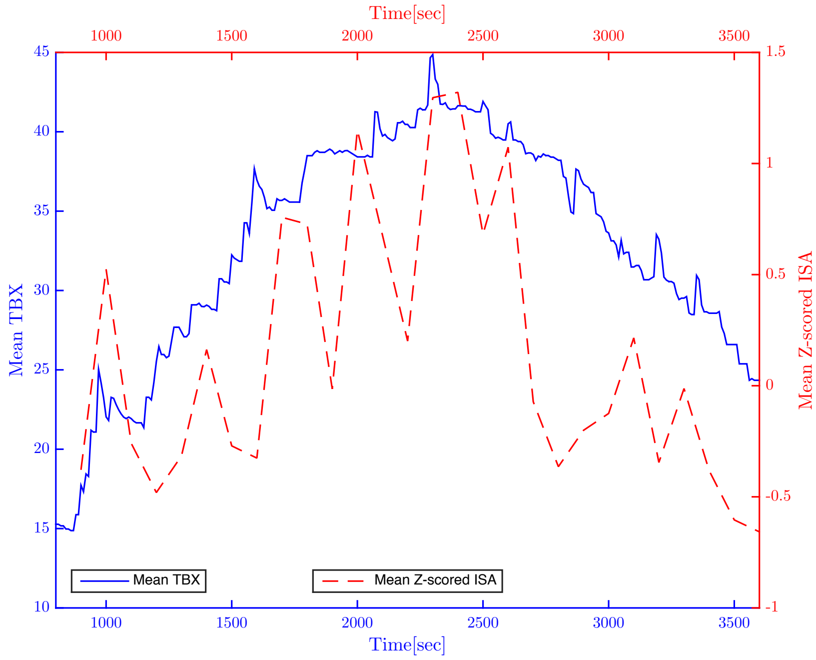


Figure H-19: Experiment time shifted 800 seconds, Z-scored mean across all participants of TBX with ISA - Unstructured Traffic, Small Perturbation

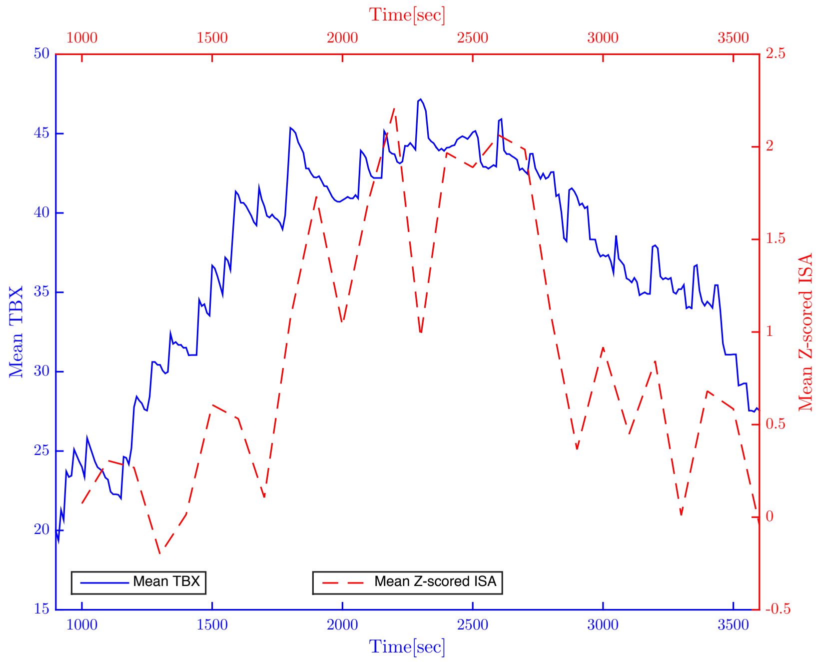


Figure H-20: Experiment time shifted 900 seconds, Z-scored mean across all participants of TBX with ISA - Unstructured Traffic, Large Perturbation

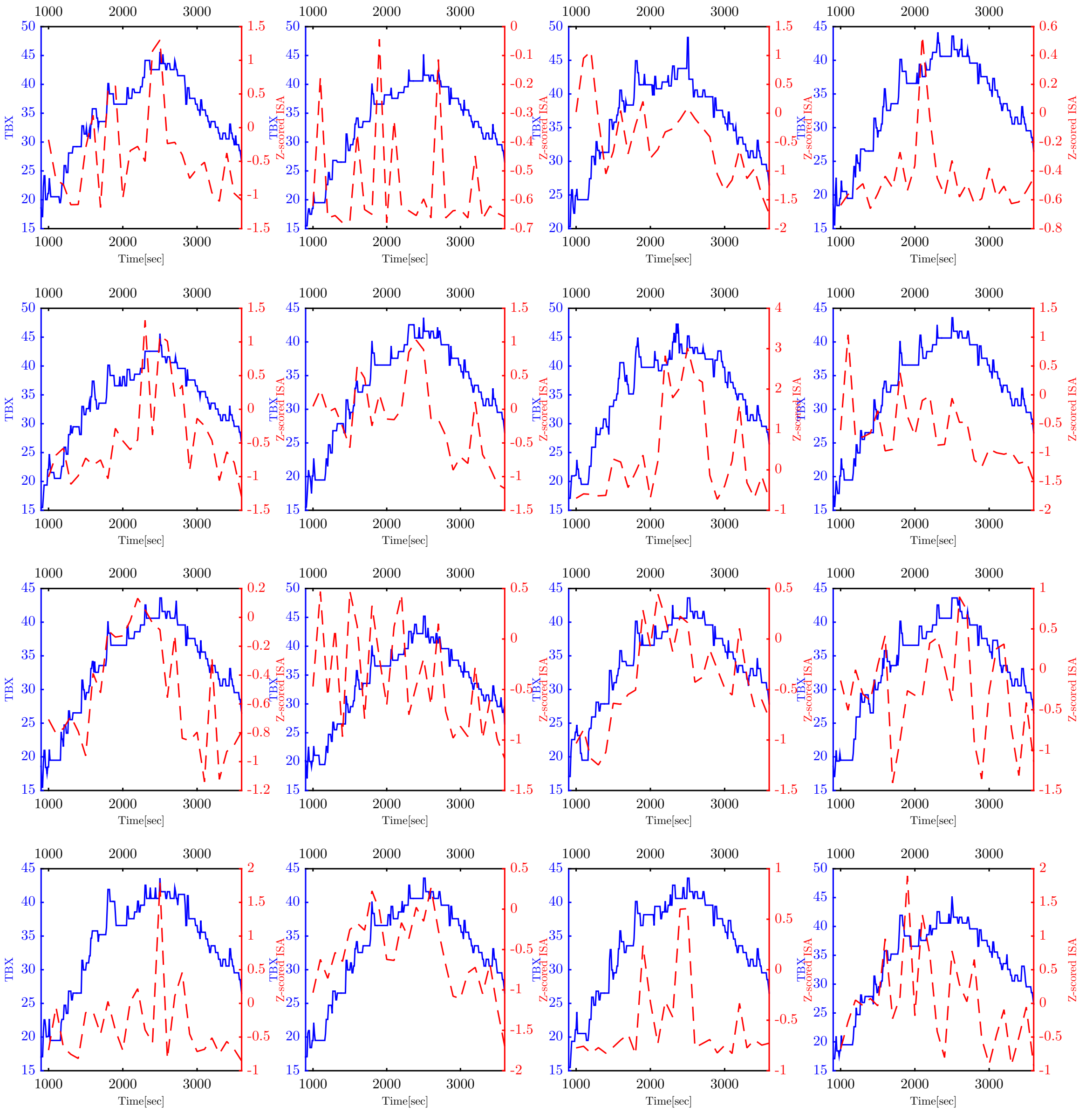


Figure H-21: All individual participant results time shifted 900 seconds of TBX with ISA - Structured Traffic, Small Perturbation

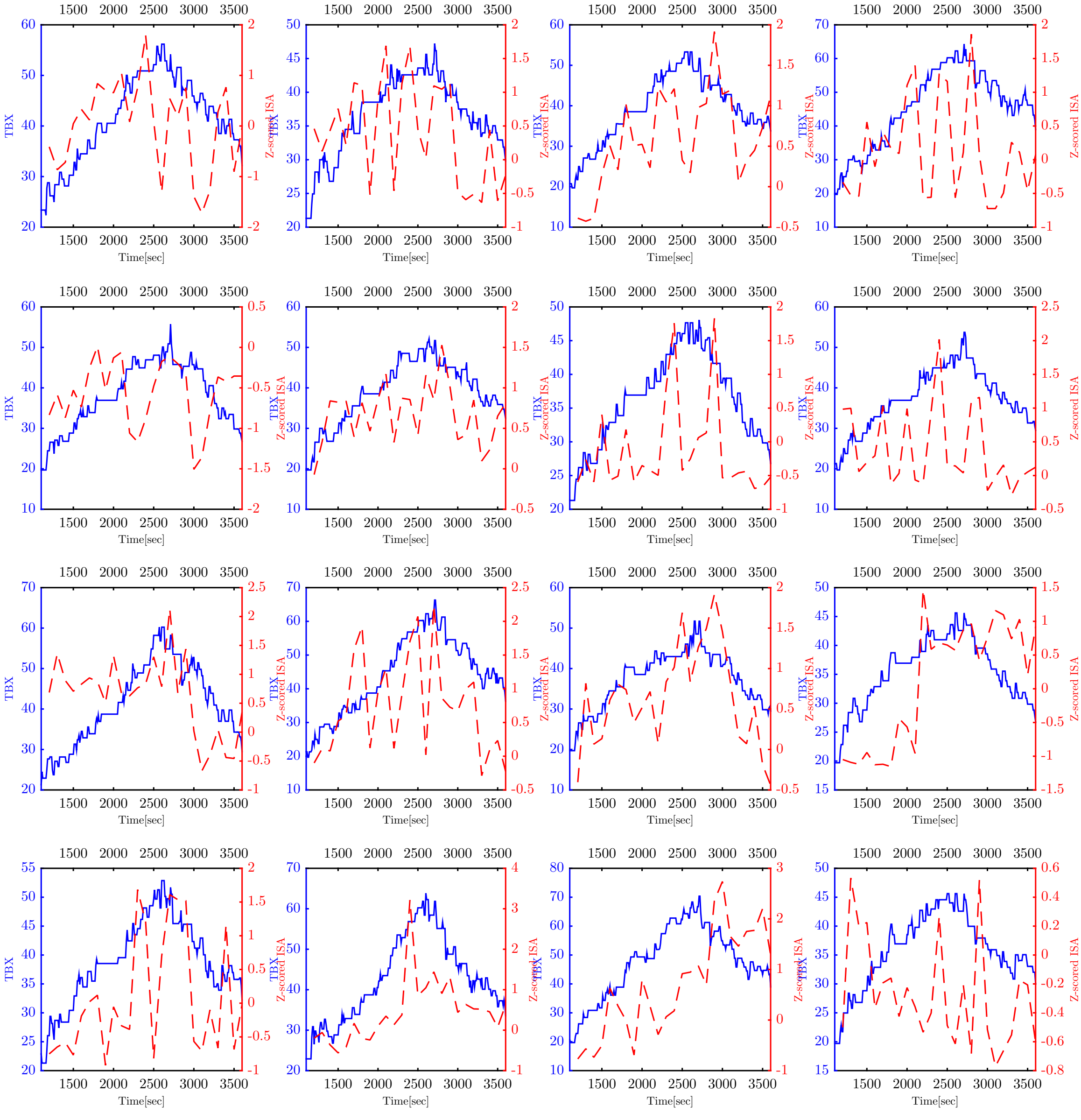


Figure H-22: All individual participant results time shifted 1100 seconds of TBX with ISA - Structured Traffic, Large Perturbation

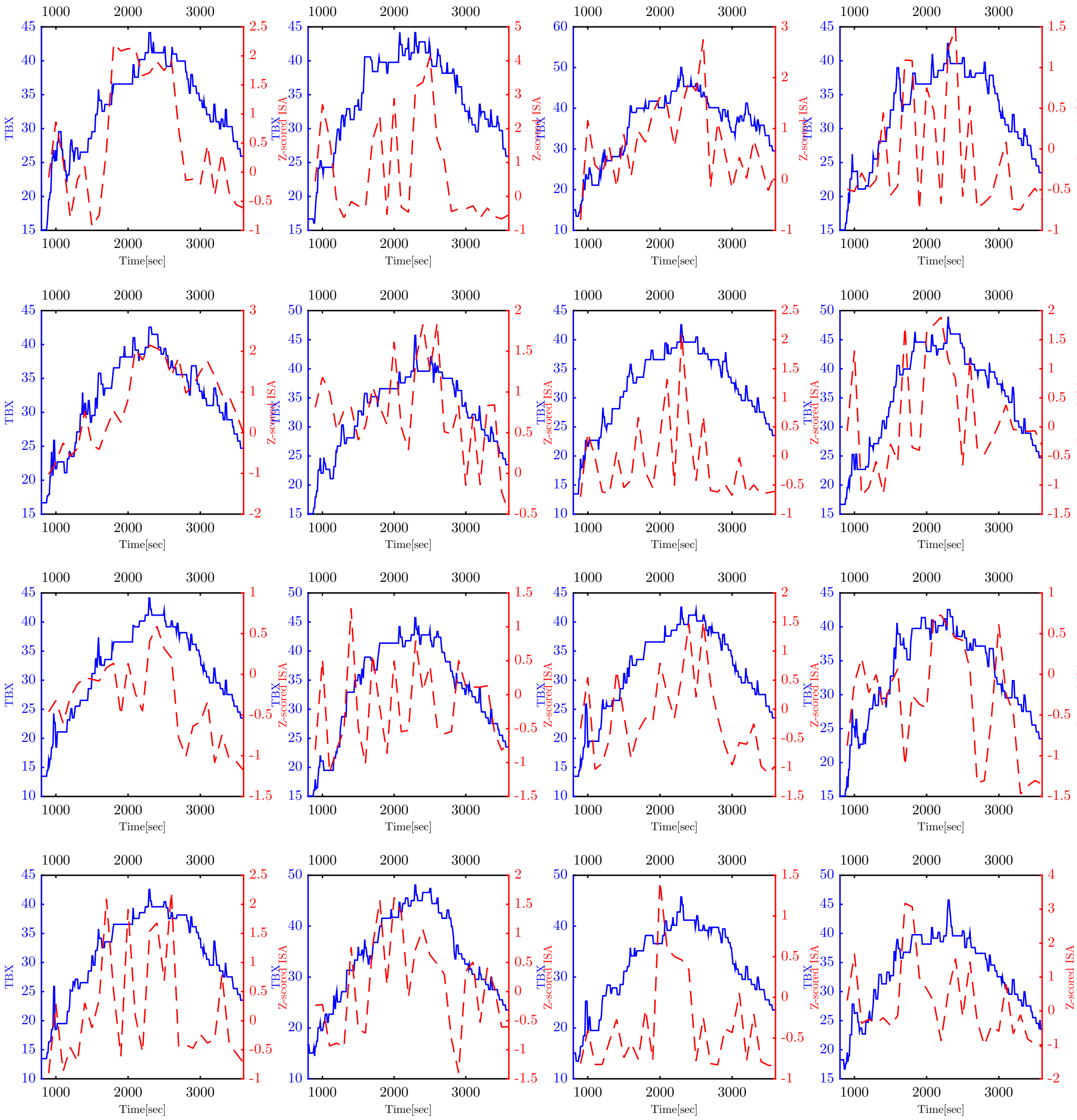


Figure H-23: All individual participant results time shifted 800 seconds of TBX with ISA - Unstructured Traffic, Small Perturbation

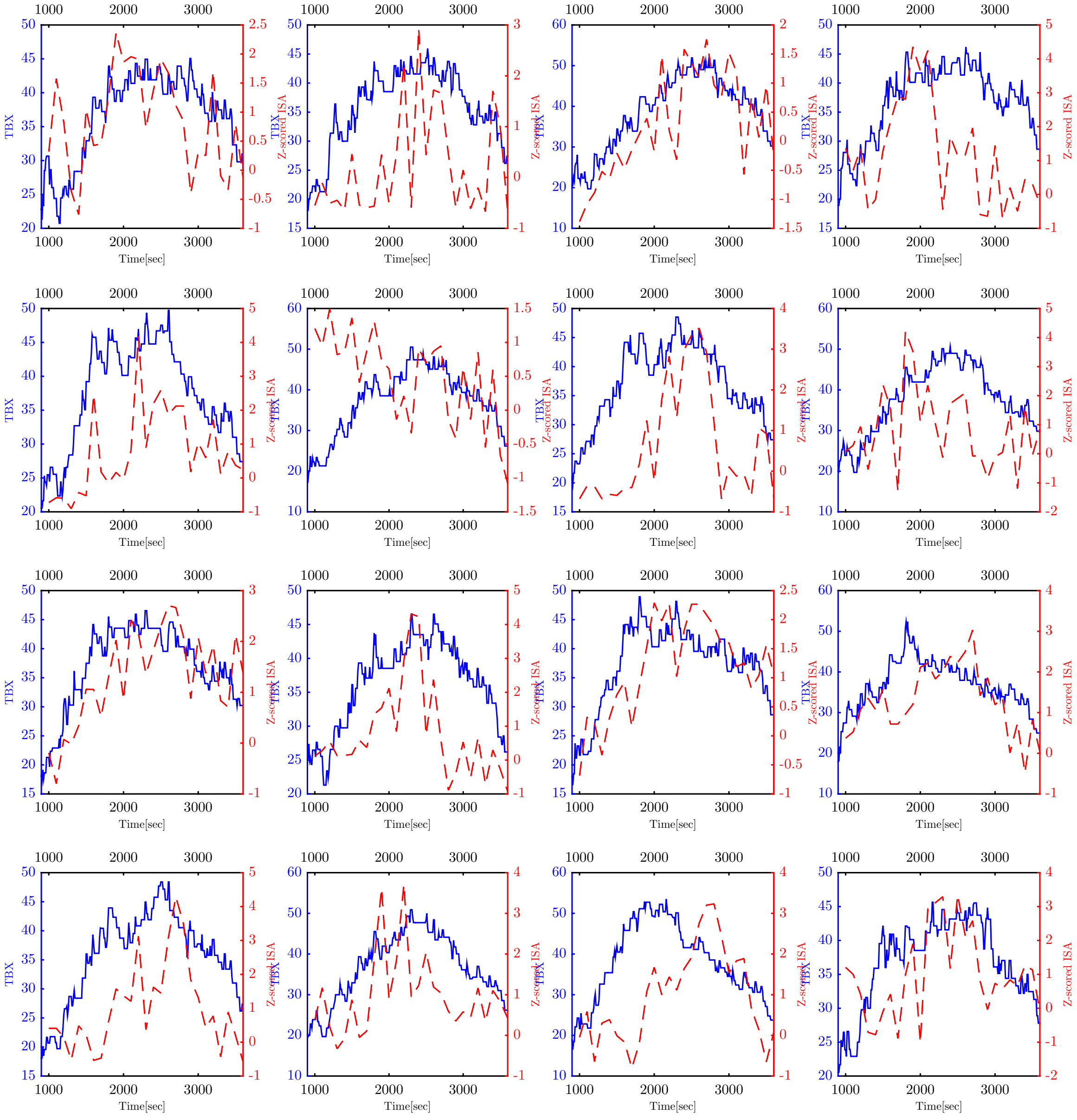


Figure H-24: All individual participant results time shifted 900 seconds of TBX with ISA - Unstructured Traffic, Large Perturbation

Appendix I

Experiment Study Results of Correlations

I-1 Dynamic Density vs Trajectory-based Complexity Correlations

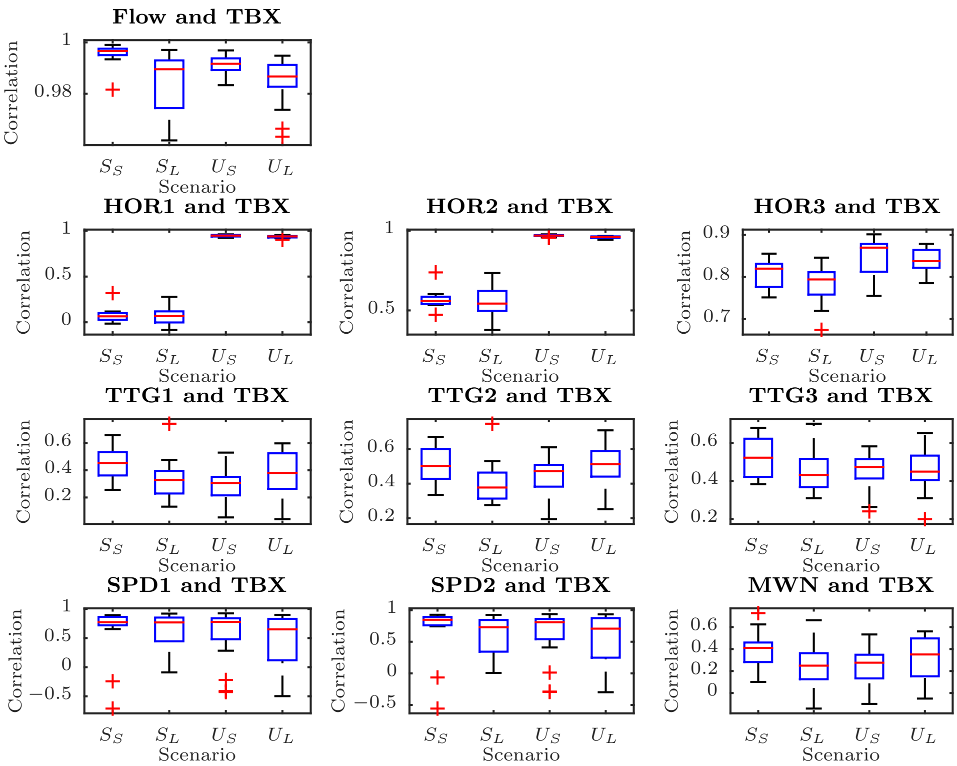


Figure I-1: Experiment Correlations of DD elements with Full TBX

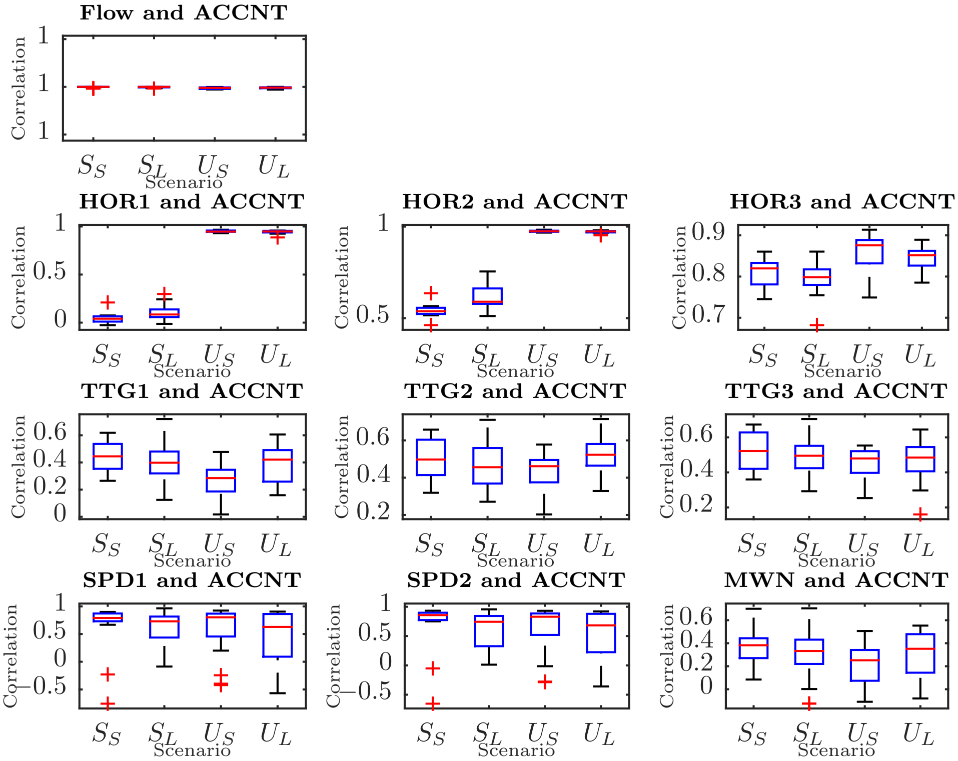


Figure I-2: Experiment Correlations of DD elements with TBX Aircraft Count

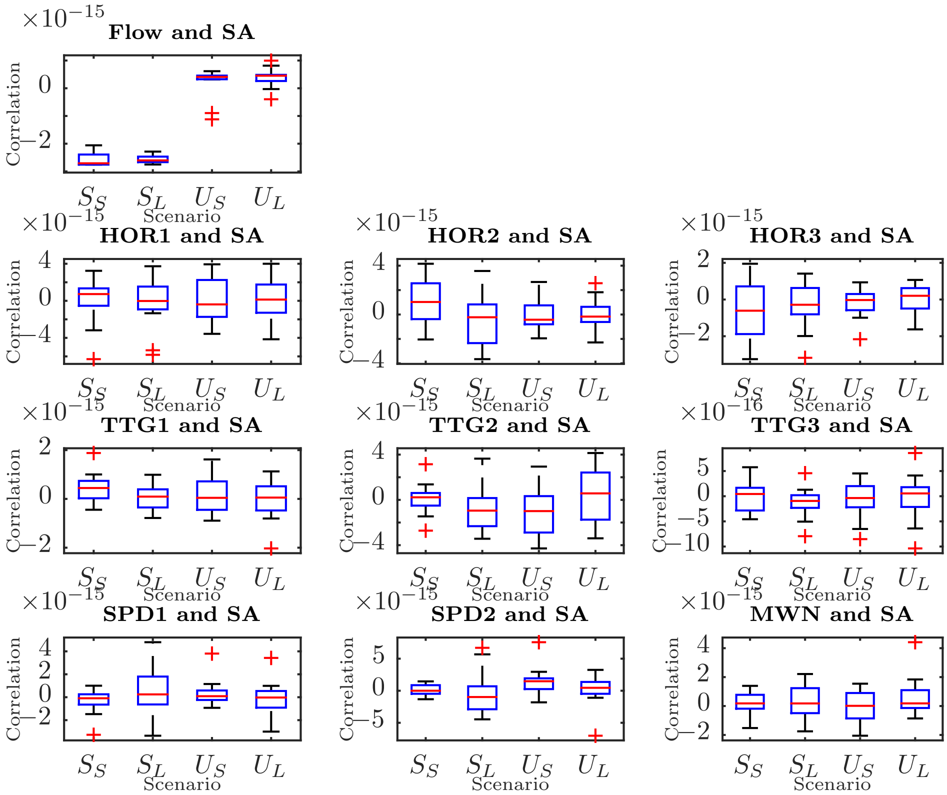


Figure I-3: Experiment Correlations of DD elements with TBX Sector Area

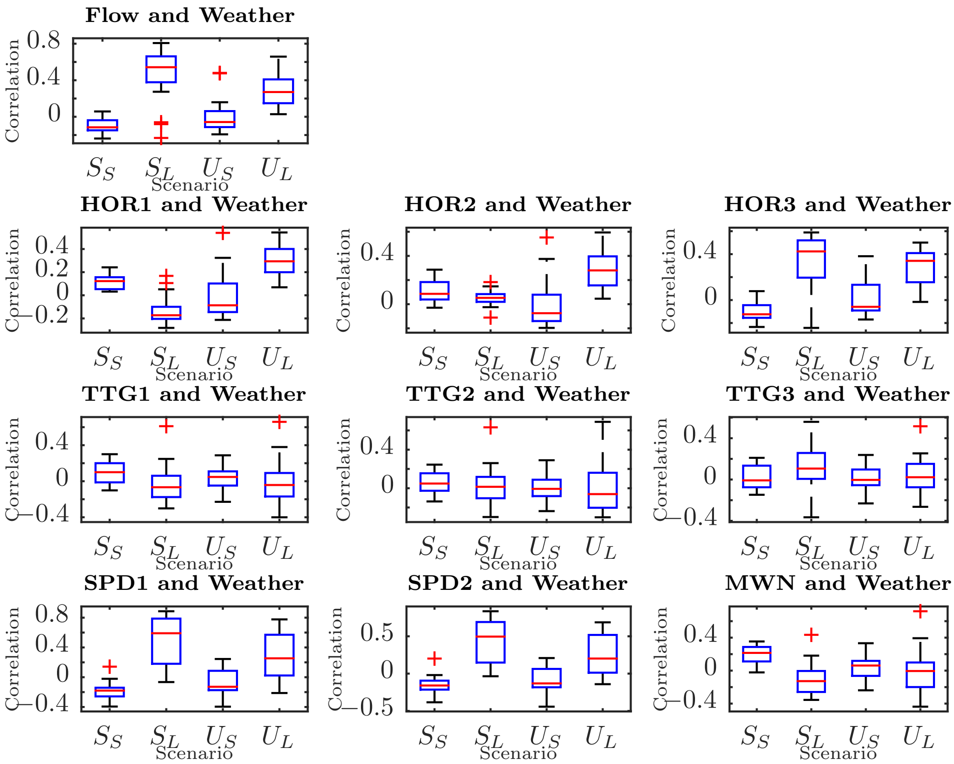


Figure I-4: Experiment Correlations of DD elements with TBX Aircraft Predicted to Penetrate Weather

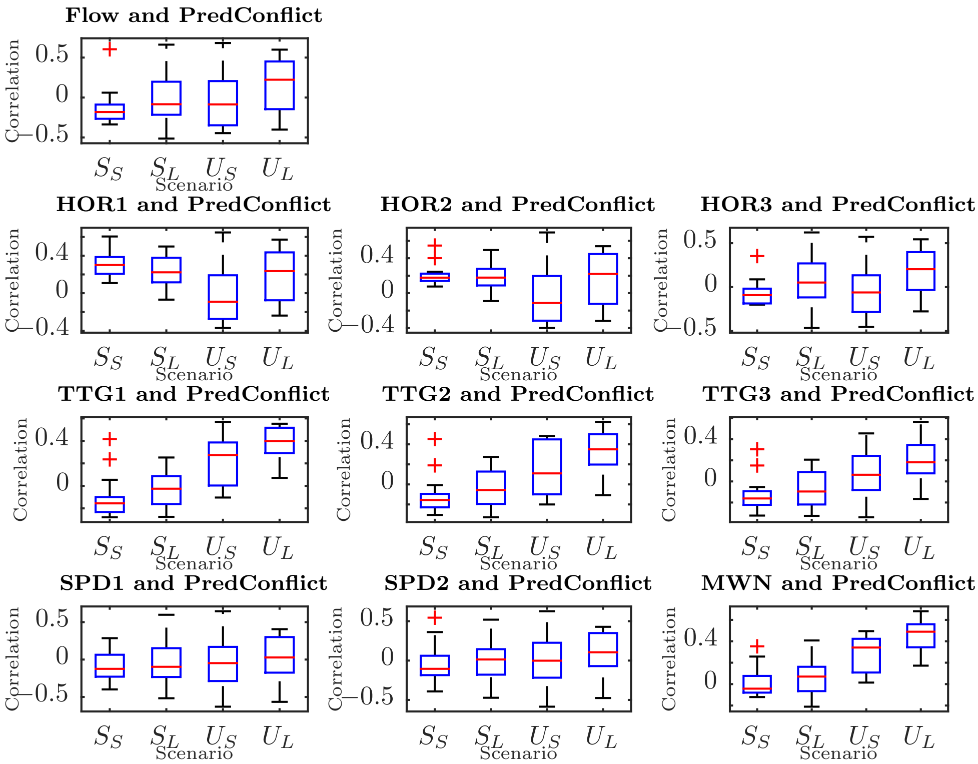


Figure I-5: Experiment Correlations of DD elements with TBX Aircraft with Predicted Conflict(LOS)

I-2 Dynamic Density vs Instantaneous Self Rating Correlations

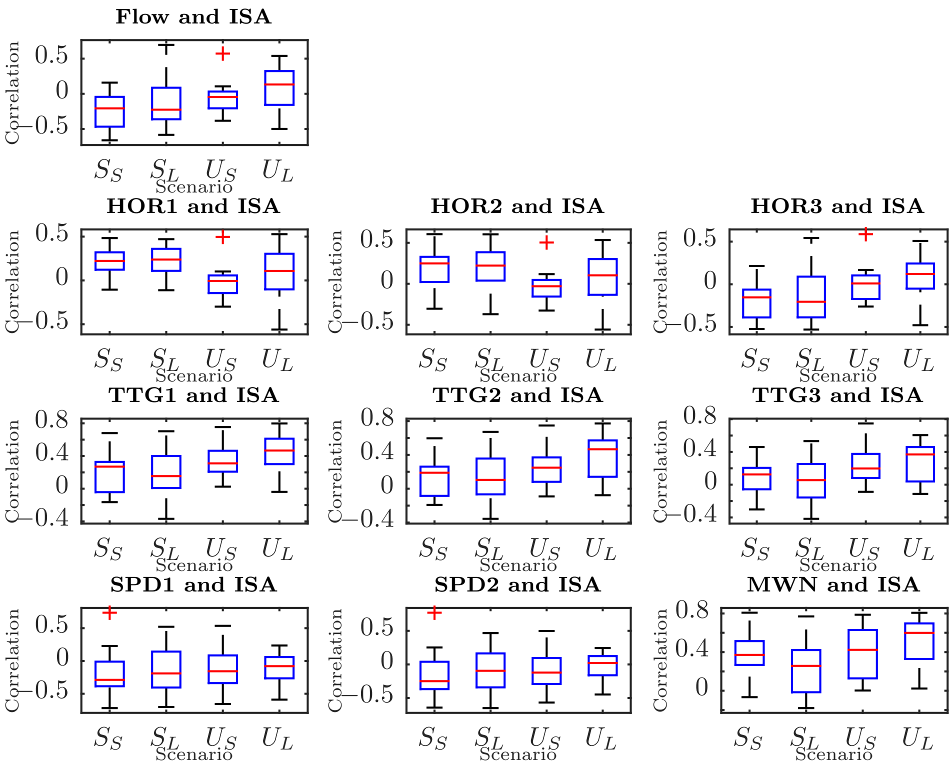


Figure I-6: Experiment Correlations of DD elements with ISA scores

I-3 Trajectory-based Complexity vs Instantaneous Self Rating Correlations

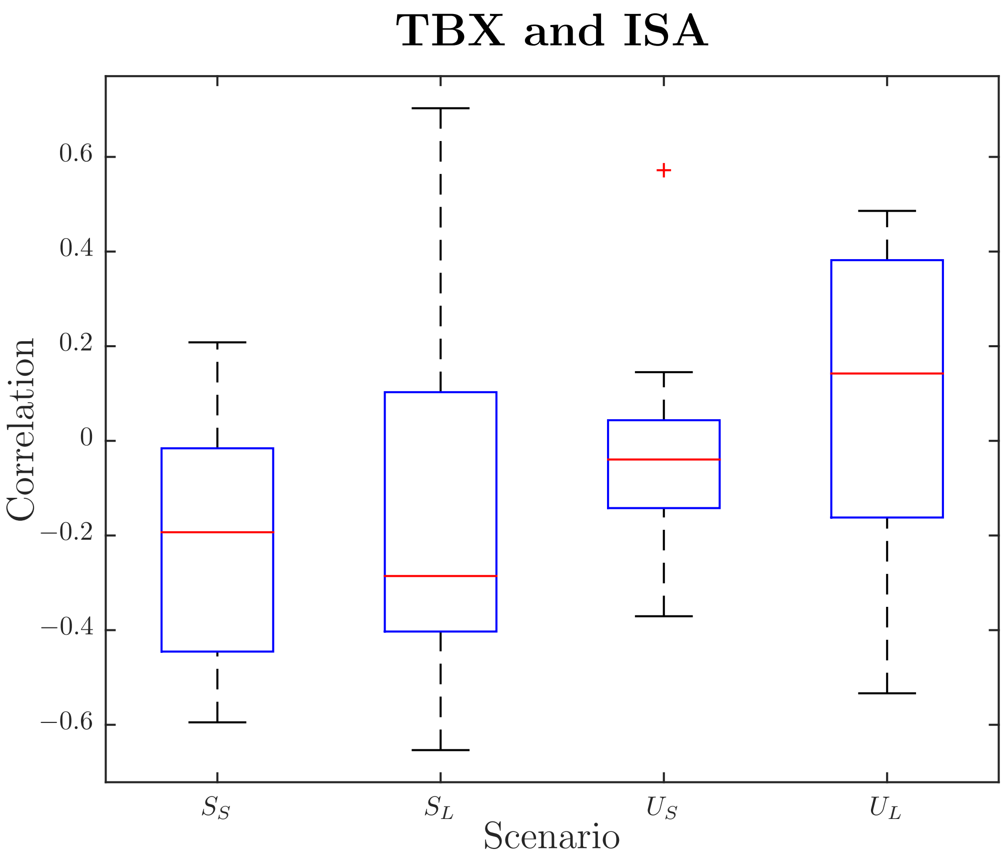


Figure I-7: Experiment Correlations of TBX with ISA scores

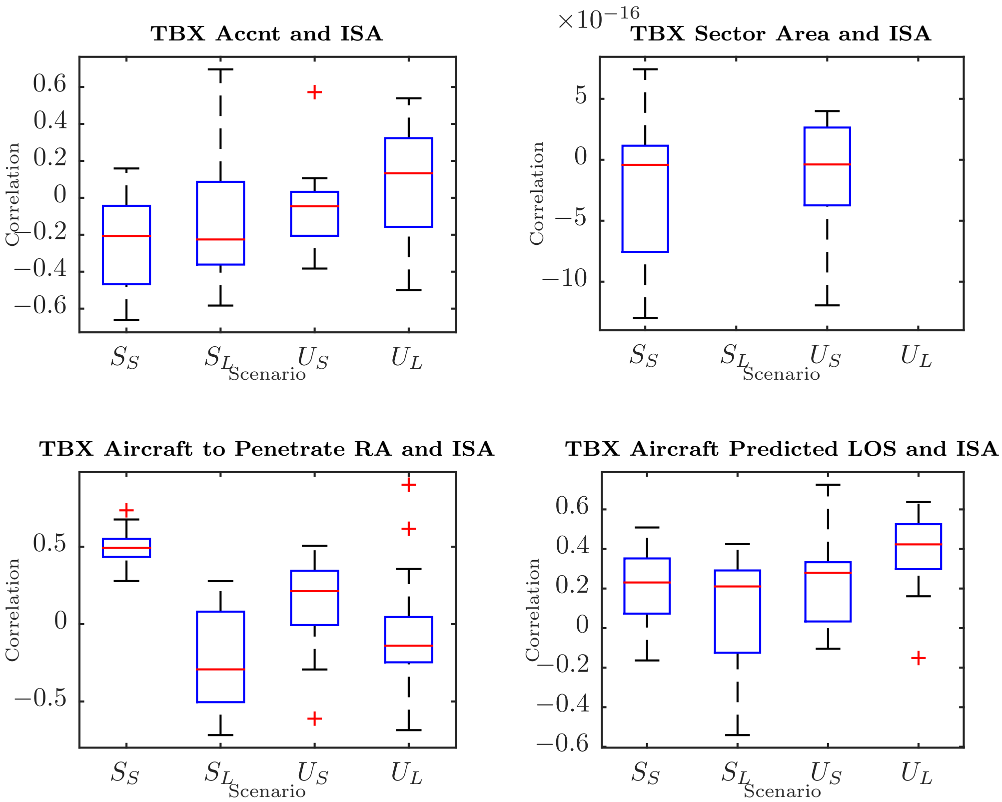


Figure I-8: Experiment Correlations of TBX elements with ISA scores

I-4 Time Shifted Trajectory-based Complexity vs Instantaneous Self Rating Correlations

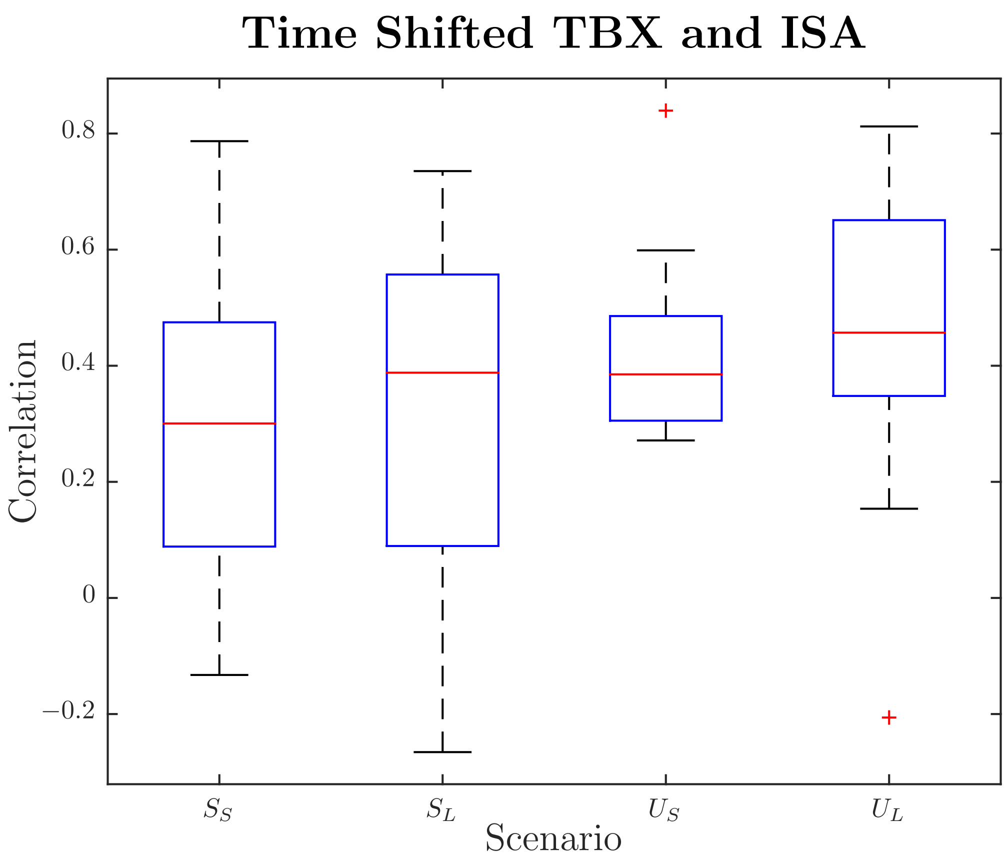


Figure I-9: Experiment Correlations of TBX with time shifted ISA scores

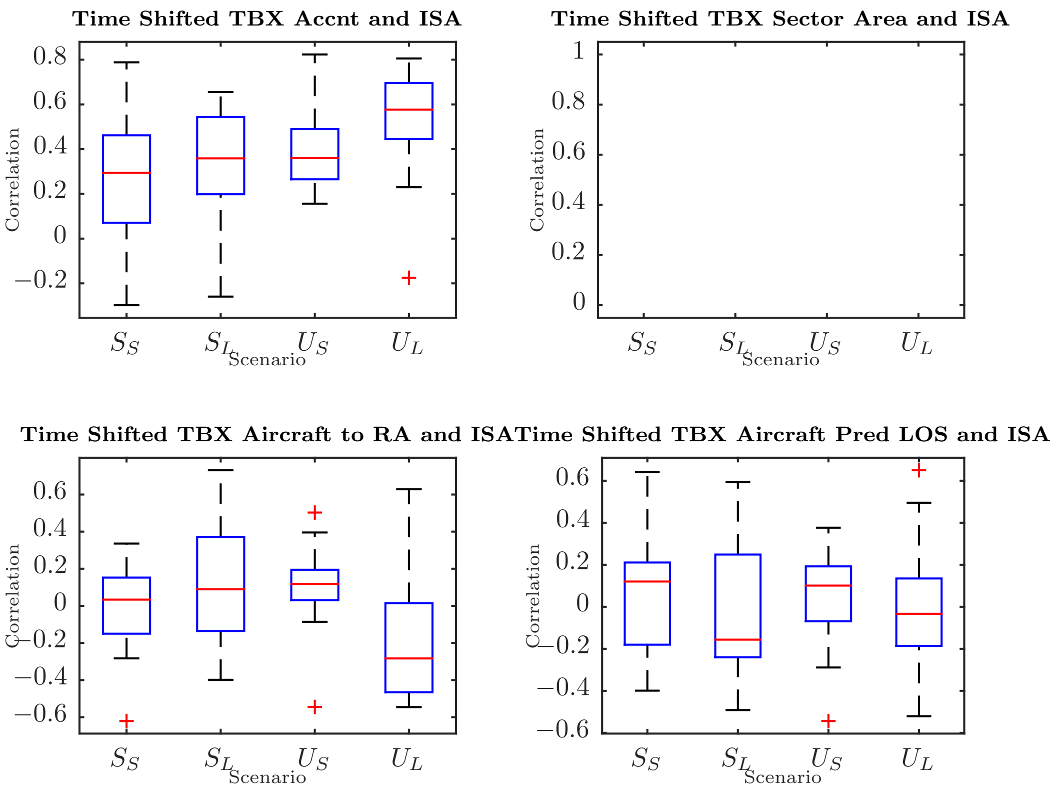


Figure I-10: Experiment Correlations of TBX elements with time shifted ISA scores

Appendix J

Past Correlations

Correlations were taken between DD complexity elements and TBX complexity elements. The data was taken from a past experiment done at TU Delft (Abdul Rahman et al., 2010) for initial investigations of the TBX complexity metric. In this section are various combinations of correlations of the elements of the two complexity metrics. Note that the values of DD are the raw values; no sector specific weightings.

J-1 Dynamic Density vs Trajectory-based Complexity Correlations

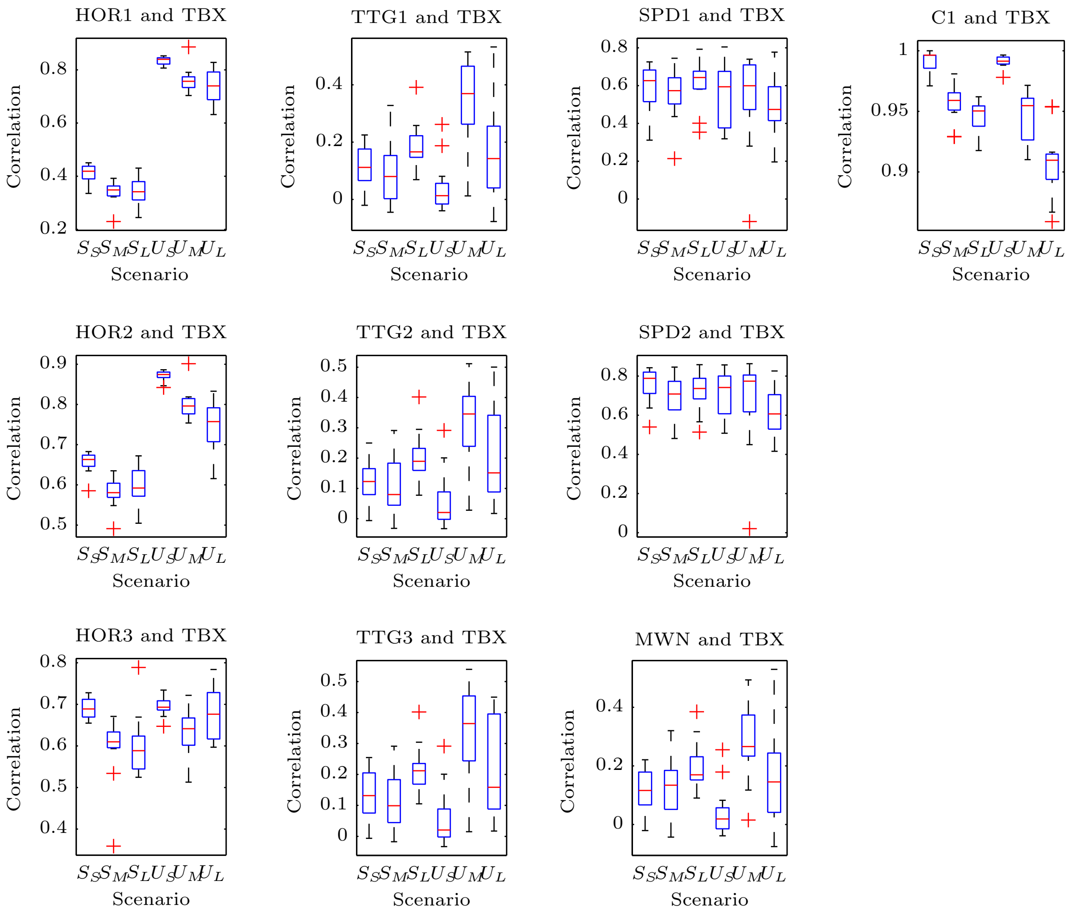


Figure J-1: Correlations of DD elements with Full TBX

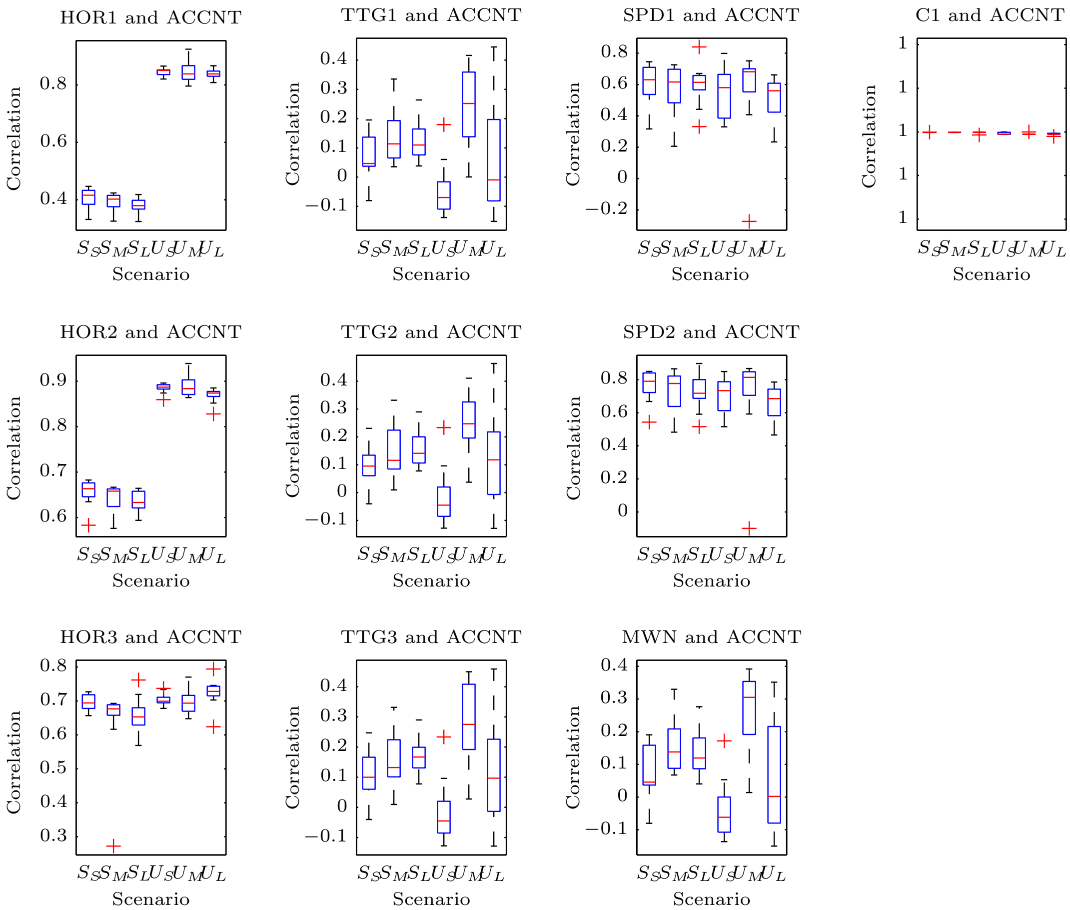


Figure J-2: Correlations of DD elements with TBX Aircraft Count

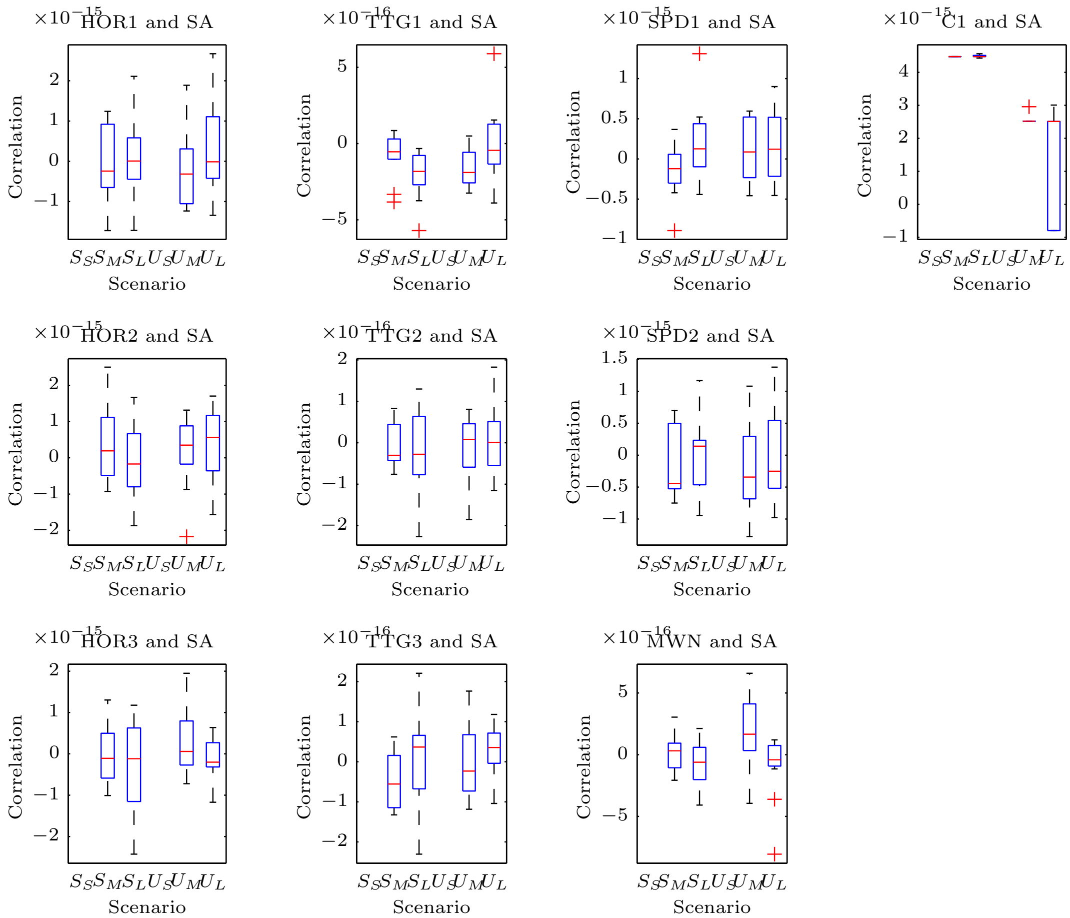


Figure J-3: Correlations of DD elements with TBX Sector Area

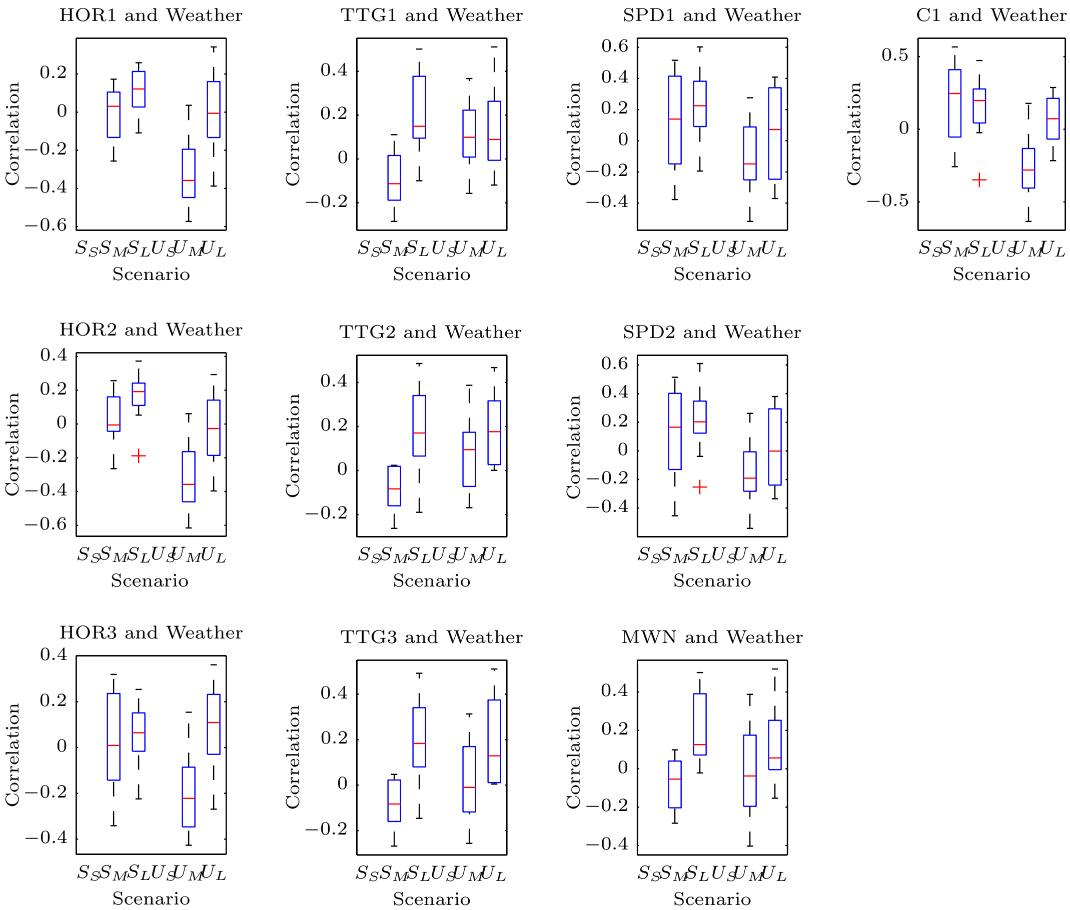


Figure J-4: Correlations of DD elements with TBX Aircraft Predicted to Penetrate Weather

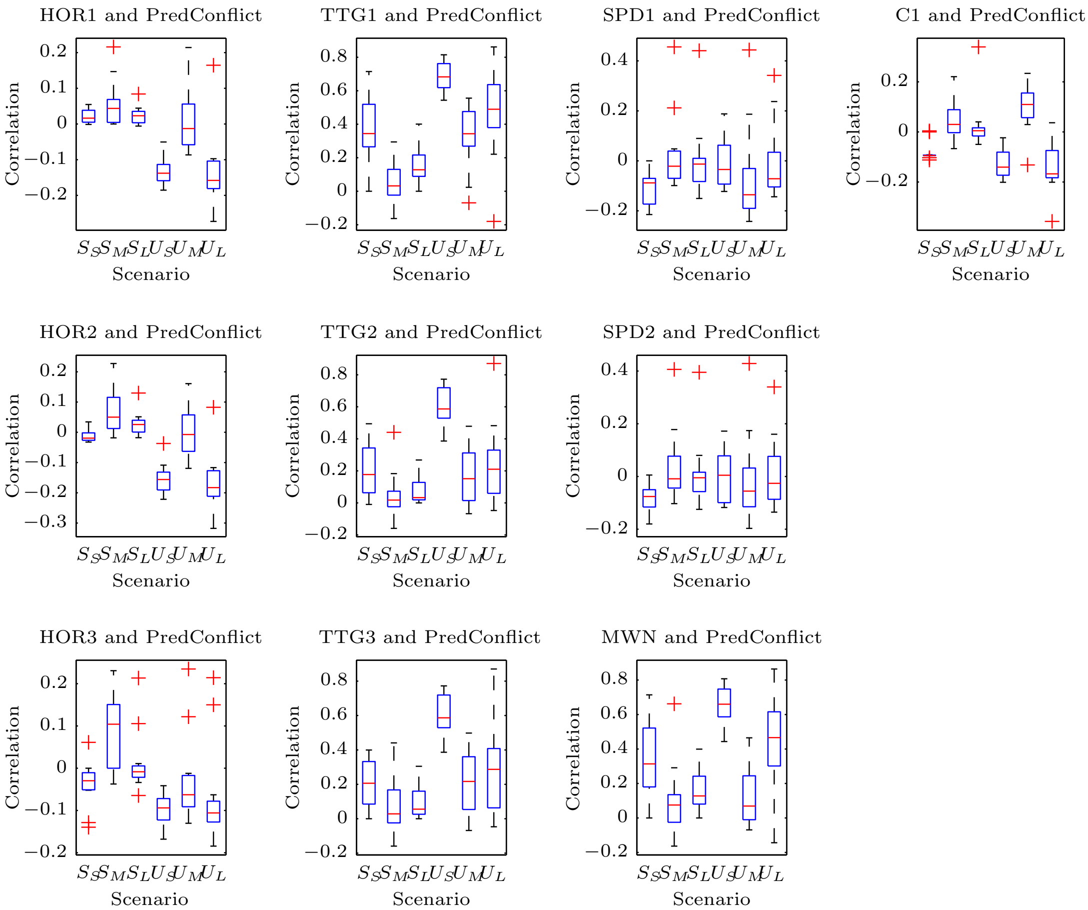


Figure J-5: Correlations of DD elements with TBX Aircraft with Predicted Conflict(LOS)

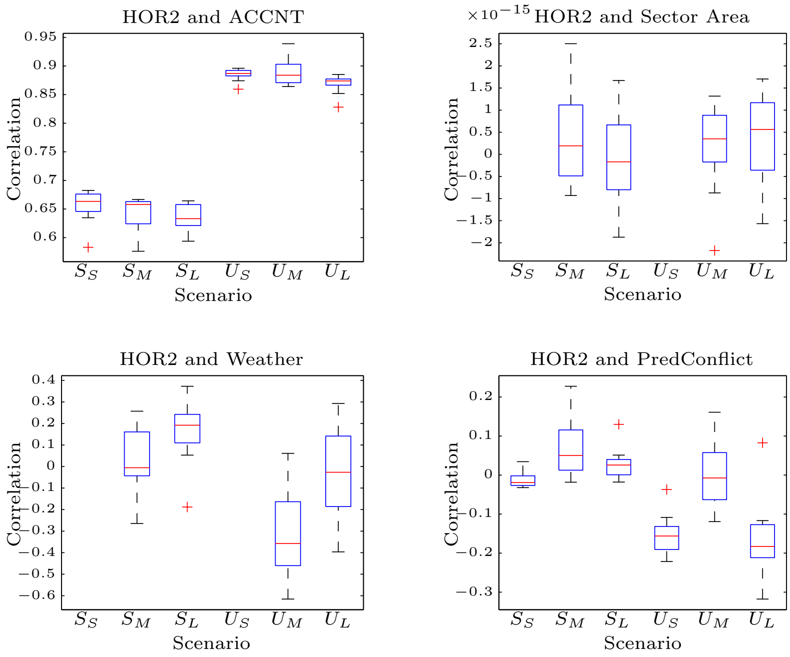


Figure J-6: Correlations of DD HOR2 with TBX elements

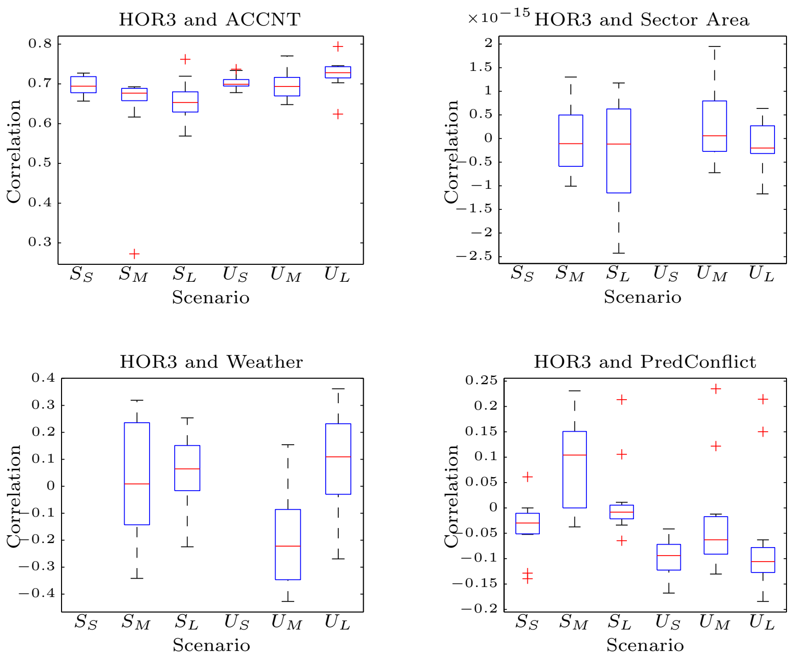


Figure J-7: Correlations of DD HOR3 with TBX elements

Appendix K

Past Complexity Metrics Plots

The plots that show the most profound trends are shown in this section as there were numerous plots generated that would not provide anymore usefulness if shown here. The plots shown are the raw DD values of HOR1 and HOR2 elements that are plotted against the full TBX value. These plots are of scenario conditions 1 and 4; structured traffic, low perturbations and unstructured traffic, low perturbations. The plots are grouped into the three research subject groups; LVNL ATCOs, NLR domain experts, and TU Delft Ph.D. students.

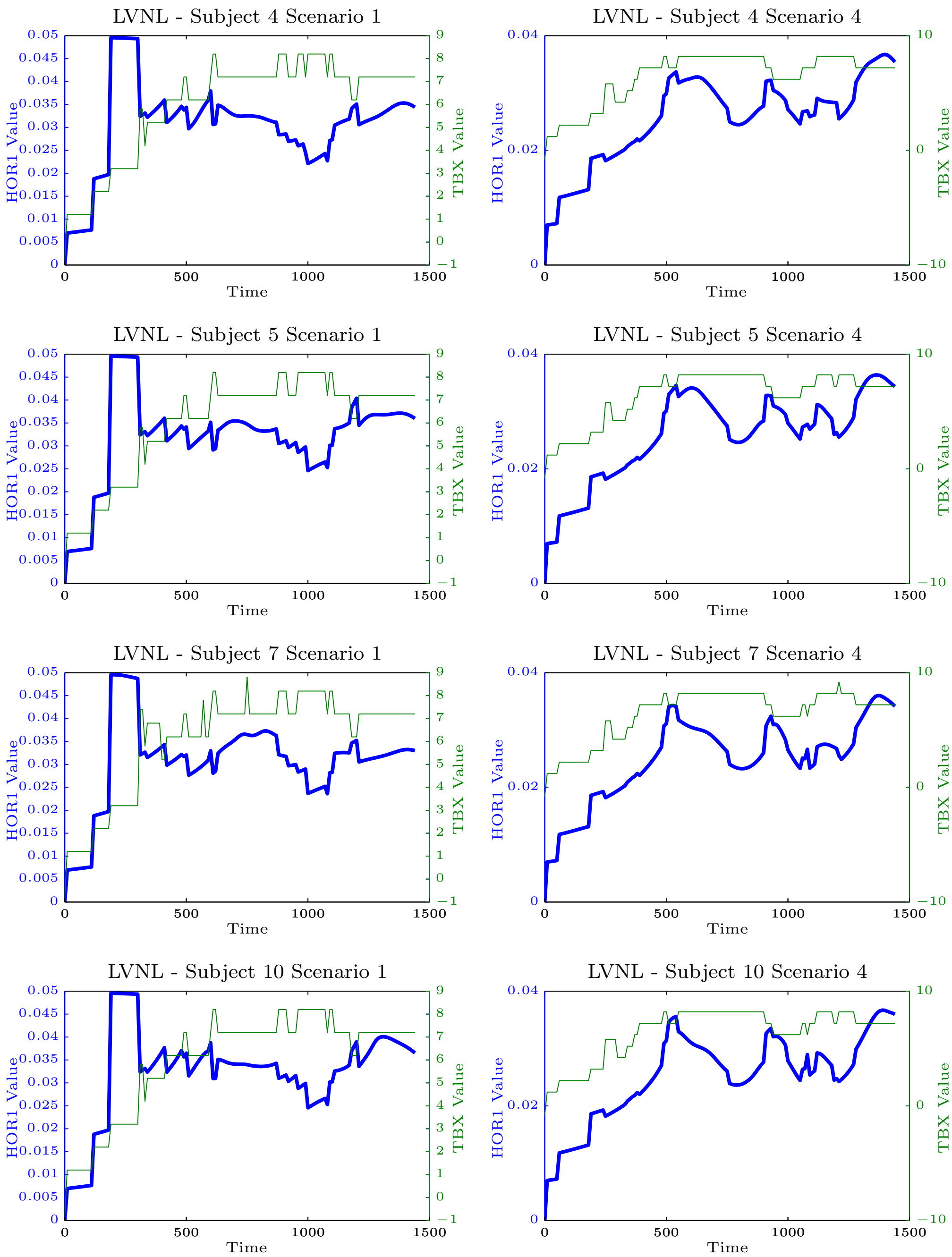


Figure K-1: ATCOs Subjects - DD HOR1 plotted against TBX

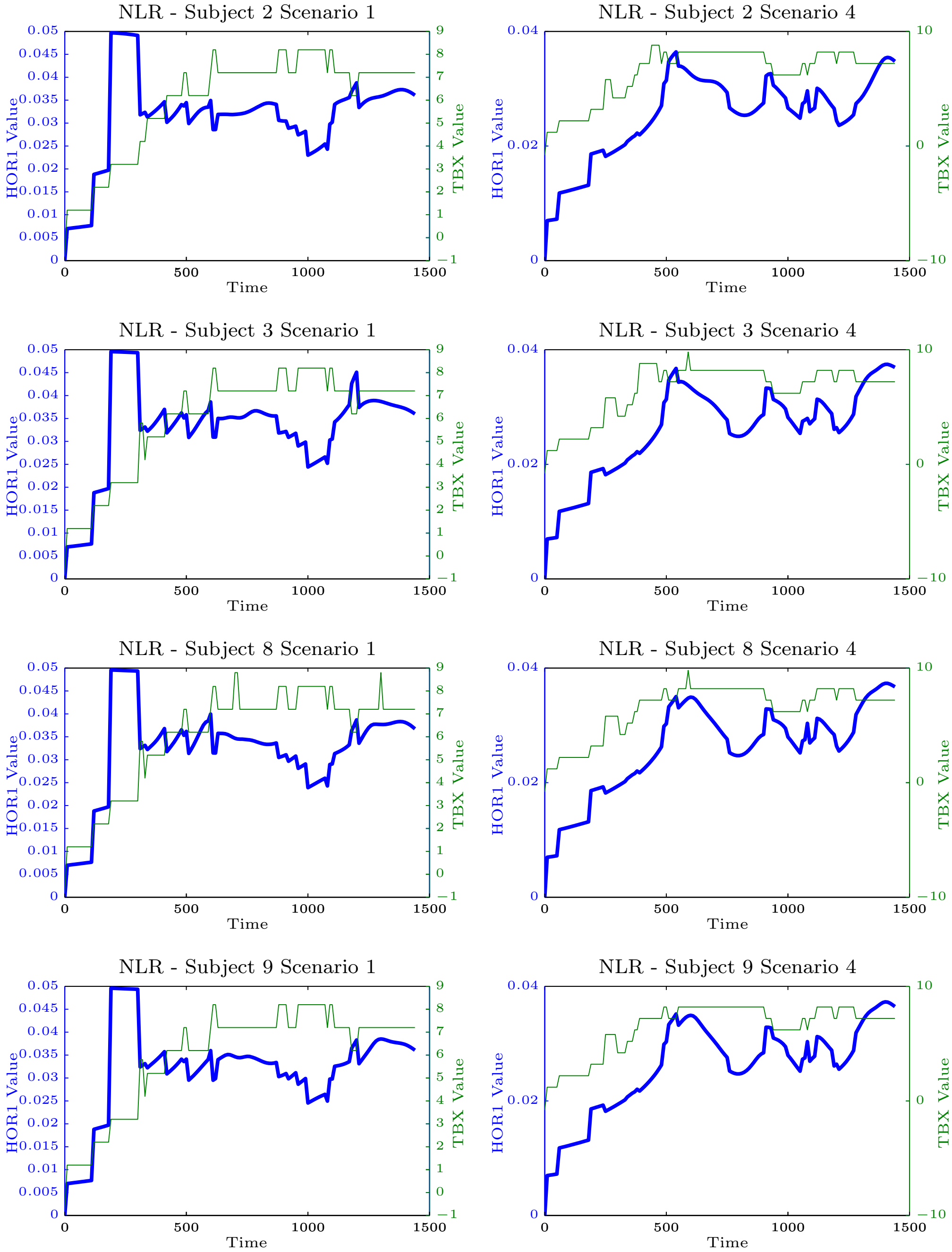


Figure K-2: NLR Subjects - DD HOR1 plotted against TBX

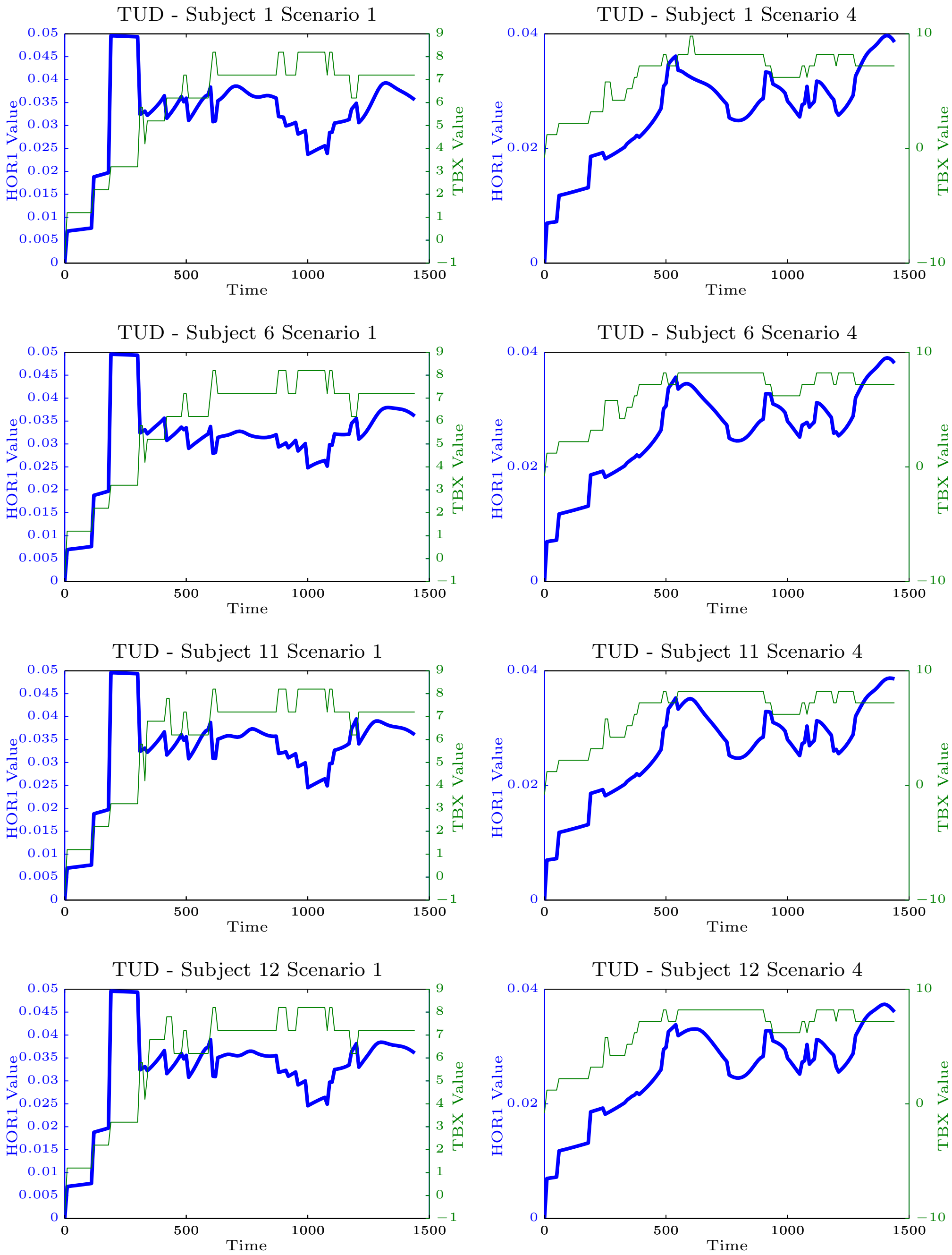


Figure K-3: TU Delft PhD Subjects - DD HOR1 plotted against TBX

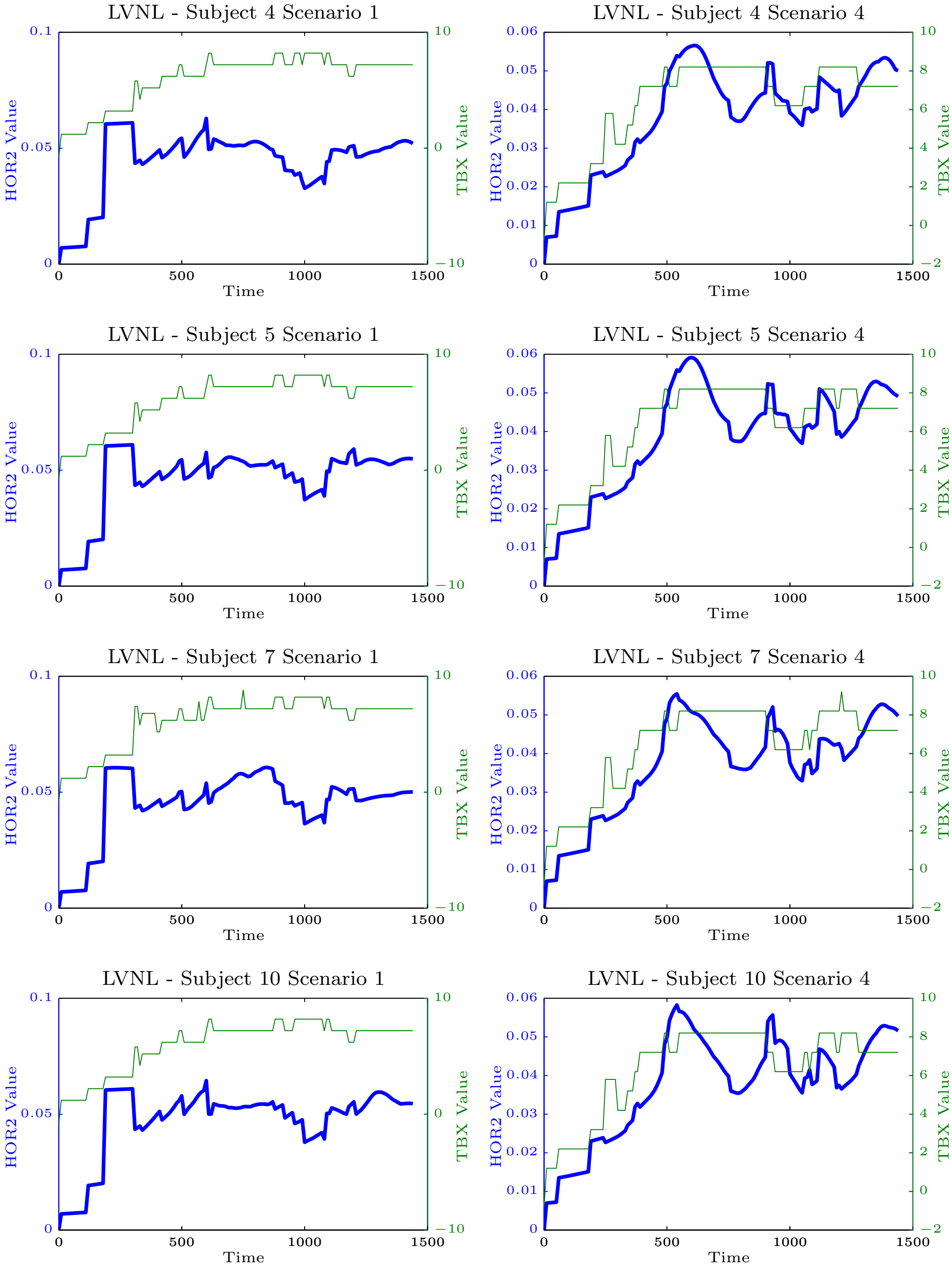


Figure K-4: ATCOs Subjects - DD HOR2 plotted against TBX

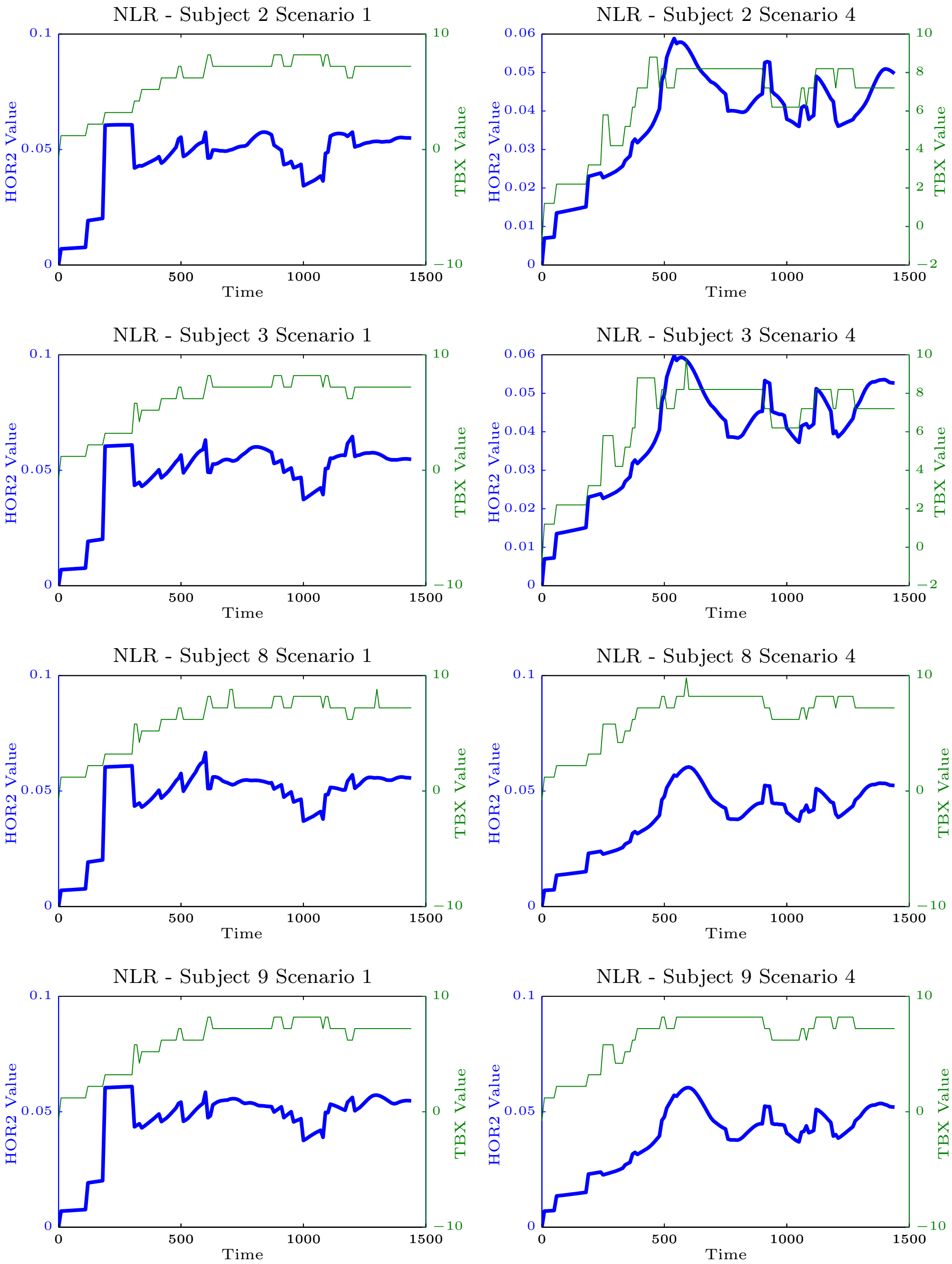


Figure K-5: NLR Subjects - DD HOR2 plotted against TBX

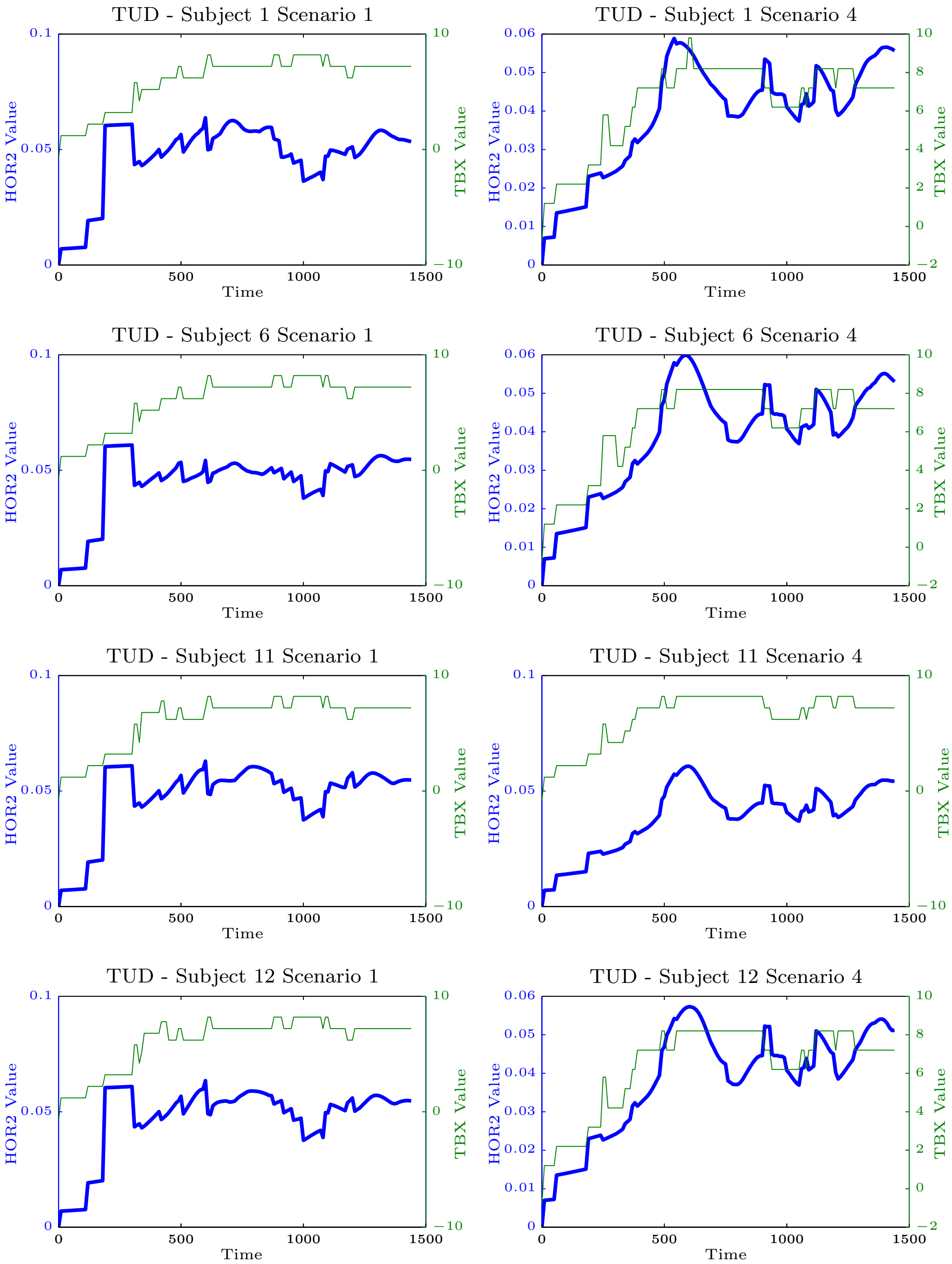


Figure K-6: TU Delft PhD Subjects - DD HOR2 plotted against TBX

Bibliography

Abdul Rahman, S. M. B., Mulder, M., & Paassen, M. M. van. (2010, August). The Effects of Air Traffic Control Sector Design on the Solution Space Diagram. In W.-P. Brinkman & M. Neerincx (Eds.), *Proceedings of ecce 2010 conference* (p. 261-268). Mediamatica, Delft University of Technology, The Netherlands.

Bainbridge, L. (1983, November). Ironies of Automation. *Automatica*, *19*(6), 775–779. Available from http://linkinghub.elsevier.com/retrieve/pii/0005109883900468

Baxter, G., Rooksby, J., Wang, Y., & Khajeh-Hosseini, A. (2012). The ironies of automation... still going strong at 30? In *Proceedings of the 30th european conference on cognitive ergonomics - ecce '12* (pp. 65–71). New York, New York, USA: ACM Press. Available from http://dl.acm.org/citation.cfm?doid=2448136.2448149

Chatterji, G. B., & Sridhar, B. (2001, October). Measures For Air Traffic Controller Workload Prediction. *Proceedings of the 1st AIAA, Aircraft, Technology Integration, and Operations Forum.* Available from http://arc.aiaa.org/doi/pdf/10.2514/6.2001-5242

Christoffersen, K., & Woods, D. D. (2002). How to Make Automated Systems Team Players. In *Advances in human performance and cognitive engineering research* (Vol. 2, pp. 1–12). Elsevier Science Ltd. Available from http://www.emeraldinsight.com/journals.htm?articleid=1781588\&show=abstract

Federal Aviation Administration. (2013, April 9). *What is NextGen?* Available from http://www.faa.gov/nextgen/slides/?slide=1

Flach, J. M., Tanabe, F., Monta, K., Vicente, K. J., & Rasmussen, J. (1998, October). An Ecological Approach to Interface Design. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, *42*(3), 295–299. Available from http://pro.sagepub.com/lookup/doi/10.1177/154193129804200324

Hilburn, B. (2004, February). *Cognitive Complexity in Air Traffic Control: A Literature Review* (Tech. Rep.). Center for Human Performance Research. Available from http://www.chpr.nl/index_htm_files/CXLIT.pdf

Kopardekar, P., & Magyarits, S. (2002). Dynamic Density: Measuring and Predicting Sector Complexity [ATC]. In *Proceedings. the 21st digital avionics systems conference.* IEEE.

Norman, D. A. (1990, April). The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, *327*(1241), 585–93. Available from `http://www.ncbi.nlm.nih.gov/pubmed/1970904http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.1990.0101`

Press, O. U. (2013, June 4). *Definition of complexity in English.*

Prevot, T., & Lee, P. U. (2011, October). Trajectory-Based Complexity (TBX): A Modified Aircraft Count to Predict Sector Complexity During Trajectory-Based Operations. In *2011 ieee/aiaa 30th digital avionics systems conference* (pp. 3A3–1–3A3–15). Seattle, Washington, USA: IEEE. Available from `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6096045`

Price, H. J. (2011, June 4). *FAA Forecast Predicts Air Travel to Double in Two Decades.* Available from `http://www.faa.gov/news/press\_releases/news\_story.cfm?newsId=12439` (Accessed 4 June 2013)

Single European Sky ATM Research. (2009, April 9). *SESAR - About Us.* Available from `http://www.sesarju.eu/about` (Accessed 9 April 2013)

Vicente, K. J., & Rasmussen, J. (1990, September). The Ecology of Human-Machine Systems II: Mediating 'Direct Perception' in Complex Work Domains. *Ecological Psychology*, *2*(3), 207–249.