Delft University of Technology

Ecological Interface Design
Sensor Failure Diagnosis in Air Traffic Control

Bijsterbosch, Vincent; Borst, Clark; Mulder, Max; van Paassen, Rene

DOI
10.1016/j.ifacol.2016.10.560

Publication date
2016

Document Version
Accepted author manuscript

Published in
IFAC-PapersOnLine

Citation (APA)
Bijsterbosch, V., Borst, C., Mulder, M., & van Paassen, R. (2016). Ecological Interface Design: Sensor Failure Diagnosis in Air Traffic Control. IFAC-PapersOnLine, 49(19), 307-312. https://doi.org/10.1016/j.ifacol.2016.10.560

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems
Aug. 30 - Sept. 2, 2016. Kyoto, Japan

Ecological Interface Design: Sensor Failure Diagnosis in Air Traffic Control

V. A. Bijsterbosch [*] **C. Borst** [*] **M. Mulder** [*]
M. M. van Paassen [*]

[*] *Control & Simulation, Faculty of Aerospace Engineering, Delft University of Technology, Delft, 2629 HS, The Netherlands (correspondence e-mail: c.borst@tudelft.nl)*

Abstract: Future air traffic control will have to rely on more advanced automation in order to support controllers in their job of safely controlling increased traffic volumes. A prerequisite for the success of such automation is that the underlying data driving it is reliable. Current technology, however, still warrants human supervision in coping with (data) uncertainties and consequently in judging the validity of machine decisions. In this paper the Ecological Interface Design (EID) framework is explored to assist controllers in fault diagnosis using a prototype ecological interface (called the Solution Space Diagram) for tactical conflict detection and resolution in the horizontal plane. Results from a human-in-the-loop experiment with sixteen participants indicate that the ecological interface with explicit presentation of the means-ends relations between higher-level functional goals and lower-level physical objects (i.e., aircraft) enables improved sensor failure detection. Especially in high complexity scenarios, this feature had a positive impact on failure detection performance.

Keywords: Ecological Interface Design, Air Traffic Control, Automation, Supervisory Control, Sensor Failure, Decision-Making.

---

1. INTRODUCTION

Predicted air traffic growth, coupled with economic and environmental realities, force the future Air Traffic Management (ATM) system to become more optimized and strategic in nature. One important aspect of this modernization is the utilization of digital datalinks between airborne and ground systems (e.g., Automatic Dependent Surveillance - Broadcast (ADS-B)) to facilitate the introduction of more advanced and more sophisticated automation. A prerequisite for the success of such automation, is that the underlying data driving it is reliable. However, alarming results with respect to ADS-B latencies and horizontal position accuracy indicate that broadcasted position errors could reach up to 7.5 nautical miles (Ali et al. (2013); Cedrini et al. (2010); Rekkas and Rees (2008); Zhang et al. (2011); ICAO (2013)), making tasks such as fully autonomous conflict detection and resolution (CD&R) error prone. Consequently, the human controller remains responsible for judging the validity of machine-generated decisions.

In an effort to support the human controller in such a task, the Ecological Interface Design (EID) framework is explored to make automation more transparent and hence improve the detection of sensor faults and judge the validity of automation advisories (Borst et al. (2015)). To this end, a prototype ecological interface called the Solution Space Diagram (SSD) for CD&R in ATM will be used (Borst et al. (2012)). The SSD reveals how traffic surrounding a controlled aircraft limits its solution options in heading and speed by means of velocity obstacles. Although the SSD has been studied in the context of decision-making, it has not yet been investigated in terms of sensor failure detection and the role of explicitly representing the so-called “means-ends” relationships between the aircraft plotted on the Plan View Display (PVD) and the velocity obstacles plotted in the SSD. We hypothesize that presenting these links will expedite fault diagnosis and monitor automation decisions to pending separation conflicts. Note that the topic of EID and sensor failure, and the explicit representation of relationships between display features, has been studied before in process control (e.g., St-Cyr et al. (2013); Burns (2000)), but not yet in the context of aviation featuring fast dynamics and short time constants.

This paper is structured as follows. First, the SSD will be briefly explained, followed by the experimental design of the human-in-the-loop study. After the results, a discussion and conclusion will be provided.

2. THE SOLUTION SPACE DIAGRAM AND THE PROPAGATION OF SENSOR FAILURES

The SSD is a constraint-based interface, designed according to the EID principles, using the state of an aircraft (internal performance constraints) and external separation criteria in terms of a Protected Zone (PZ) to indicate the solution space in terms of heading and speed. This enables controllers to detect conflicts (when the speed vector of a controlled aircraft lies inside a conflict zone) and avoid a Loss of Separation (LoS) by giving heading and/or speed clearances to aircraft in order to direct the speed vector outside a conflict zone (Fig. 1).

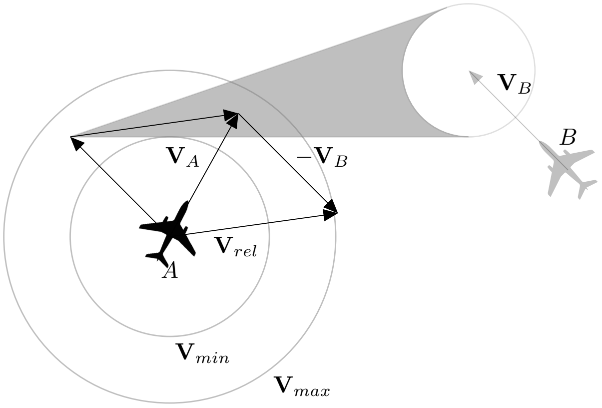
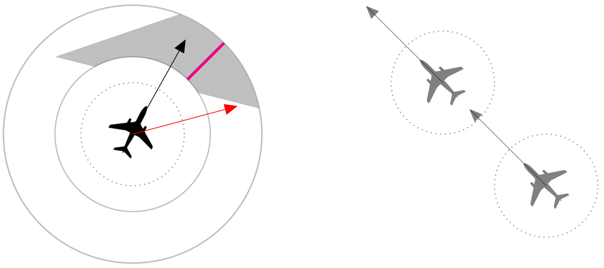


Fig. 1. The SSD, showing the triangular conflict zone (imposed by aircraft B) within the speed envelope of aircraft A.

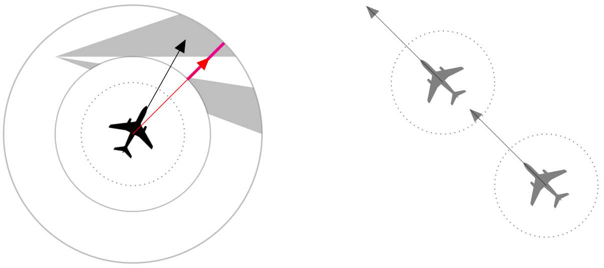
The shape and orientation of the conflict zone can give the controller information about the location and proximity of neighbouring aircraft. That is, the cone of the triangle points toward, at a slight offset, the neighbouring aircraft and the width of the triangle is large for near-by aircraft and small for far-away aircraft. Additionally, drawing an imaginary line from the aircraft blip toward the tip of the triangle indicates the absolute speed vector of a neighbouring aircraft. As such, with the shape and orientation of the conflict zones, a controller would be able to link aircraft to their corresponding conflict zones.

To construct the SSD, however, detailed information about the position and velocity vectors of aircraft need to be available, for example through ADS-B. It is also very likely that in the transition phase toward using ADS-B as a primary means of surveillance, position information will remain available from primary and/or secondary surveillance radar given the inaccuracies in current ADS-B systems. This implies that discrepancies between ADS-B and the radar image may arise, resulting in an ambiguity between the aircraft position shown on the PVD (source: surveillance radar) and the representation of the conflict zone (source: ADS-B).

Additionally, CD&R automation may generate advisories (and plot them within the SSD) based upon faulty ADS-B information, potentially masking the ambiguity and thus make it difficult for the controller to judge the validity of the given advice. In Fig. 2 an example traffic situation is shown that illustrates the ambiguity between correct aircraft positions plotted on the PVD and the conflict zones plotted in the SSD in case of ADS-B position errors. Although the advisory may appear to be correct as shown in Fig. 2(b), the conflict zone formed by the aircraft in the lower right corner does not match its image shown on the radar plot. That is, the SSD suggests the trailing aircraft is much further behind the leading aircraft. This results in an erroneous solution space between the two conflict zones, suggesting that the controlled aircraft can safely be vectored in-between the two neighbouring aircraft. Although in this situation the fault could be easily spotted, one can imagine that in more dense and complex traffic situations this error is much harder to detect.



(a) Conflict geometry and advisory without failure.



(b) Conflict geometry and advisory with failure.

Fig. 2. The effect of sensor failures on the SSD and resolution advisory.

Several studies investigating EID and sensor failure detection have shown that making the relationships between functional constraints and interface objects more explicit will improve fault diagnosis (see Burns (2000); St-Cyr et al. (2013)). In this context, making the relationships between the conflict zones in the SSD and the aircraft on the PVD more explicit should expedite fault diagnosis. But is this also true when the work domain dynamics are fast and complex, requiring swift and correct controller responses?

3. EXPERIMENT DESIGN

An experiment has been designed and conducted, to investigate whether or not explicitly representing the means-ends relations between aircraft and their conflict zones positively contributes to sensor failure detection and diagnosis, irrespective of traffic complexity.

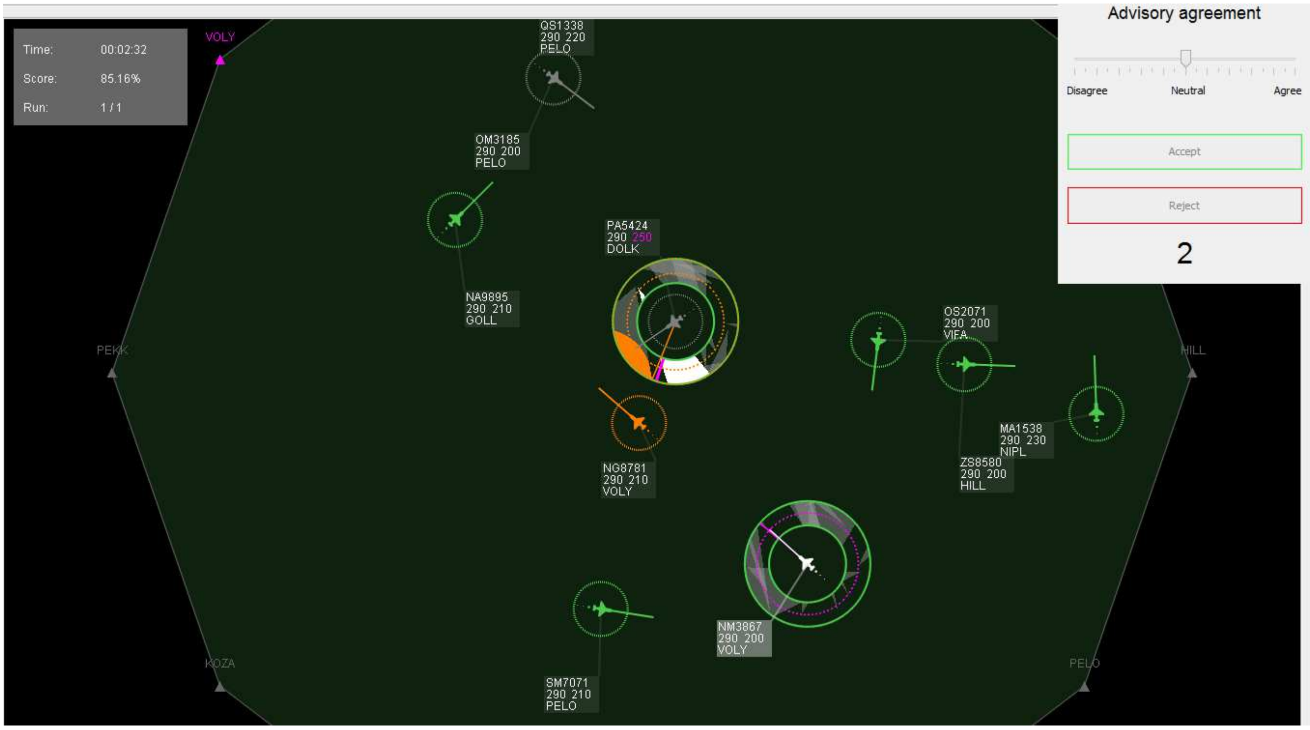
3.1 Participants and tasks

Sixteen participants volunteered in the experiment, all students or researchers in the Control & Simulation (C&S) department of Aerospace Engineering at TU Delft. Their experience varied from working in Air Traffic Control (ATC) and ATM domains to aircraft control systems.

The control task given to the participants was two-fold, namely:

(1) **Conflict resolution task:** The primary control task was to ensure safe separation (at least 5 nautical miles) between aircraft by resolving conflicts highlighted by automation.
(2) **Exit clearance task:** The secondary control task was to ensure all aircraft exit the sector airspace at their designated exit points.

To simulate a supervisory control setting featuring a high level of automation, both control tasks could be com-



Time: 00:02:32

Score: 85.16%

Run: 1 / 1

Advisory agreement

Fig. 3. Screenshot of the ATC simulator at the moment of an advisory.

pleted by either accepting or rejecting a given automated resolution advisory within 30 seconds ("management-by-exception", see Parasuraman et al. (2000)), and additionally by manually vectoring those aircraft of which the advisory has been rejected. It was stressed to the participants to carefully inspect the validity of the SSD with the overall traffic situation given by the radar image before accepting or rejecting an advisory.

In reality, the advisories were scripted, but the participants were told that the advisories were generated by the automation that checked for conflicts and exit point deviations. It was told that the advisories were based on the information from the SSD, which on its turn is generated by ADS-B data. Since this research focuses on the performance if the interface is subjected to ADS-B sensor failure, there will be cases where the SSD is inconsistent with the overall traffic situation visualized by the radar. This means that there were test cases where an advisory was given that was correct according to the SSD, but in reality was misleading and would result in a conflict. These cases specifically were of interest to investigate the sensor failure detection and diagnosis.

3.2 Apparatus

The ATC simulator is a Java-based application enabling controllers to vector aircraft in a given airspace, using the SSD tool. Figure 3 shows a screenshot of the simulator during an advisory, showing the agreement rating window on the top right and two open SSD's, one for the advisory and one manually opened by the controller. Furthermore, a conflict zone is highlighted when the mouse cursor hovers over the aircraft on the radar plot. This feature makes the means-ends link between aircraft and conflict zone more explicit.

3.3 Independent Variables

Three independent variables were defined, resulting in a mixed within- and between-subjects experiment.

First, the between-subjects variable was the explicit representation of the means-ends relations, resulting in two different interfaces tested by two separate groups (**group I:** means-ends relations off and **group II:** means-ends relations on).

Second, a within-subjects variable was the scenario complexity, resulting in two levels, namely low complexity (**Lx**) and high complexity (**Hx**). The scenario complexity is a derivative of structured versus unstructured air traffic flow. By keeping the aircraft number approximately the same at all time, the average conflict free solution space for the unstructured, high complexity situation was smaller.

Third, a within-subjects variable was the sensor failure, which can be either on **Y** or off **N**. The sensor failure was always an ADS-B position off-set of 7.5 nautical miles, which in literature study is found to be a realistic, frequently occurring error. In the scenarios with a sensor failure, only one aircraft emitted incorrect ADS-B position data. The sensor failure impacted the SSD and the automation advisory. Hence, the failures scenarios tested in the experiment were as shown in Table 1.

Table 1. The failure scenarios tested in the experiment, which exclude false positives and false negatives.

	Automation	
Sensor	correct	incorrect
correct	✔	-
incorrect	-	✔

3.4 Control Variables

The SSD version used for this research was the Time-To-Contact (TTC), meaning that the time it takes for a LoS to occur is color-coded in darkgrey (over 180 sec.), orange (between 90 and 180 sec.) and red (between 0 and 90 sec.), as can be seen in Fig. 3. These visual cues should help the controller to prioritize which conflicts are critical and need urgent attention. In order to avoid confounds, all aircraft used in the scenarios were of the same class, meaning they have the same performance envelope (velocity range) and were a fixed at a certain altitude. Furthermore, the airspace sector dimensions were kept constant for all scenarios, with shape of a decagon of 50 nautical miles radius.

In order to test multiple traffic situations per scenario, and to keep the scenarios repeatable and interesting, the simulation run at 3-times faster than real time. This resulted in a traffic scenario of 585 seconds, which ran for 195 seconds in the simulation.

3.5 Dependent Measures and Resolution Advisories

Several measurements were taken of which the most interesting are listed here.

- **Correct accept/reject score** measured if a participant agreed with the advisory or wanted to implement his/her own solution.
- **Advisory agreement rating** measured the level of agreement with the given advisory, which was measured by a slider bar with scale 0-100 before responding to the advisory.
- **Sensor failure detection** was measured using verbal comments and where it was noted whether the correct sensor failure was detected and the corresponding aircraft identified.
- **Advisory response time** measured the time between initiation of, and response to advisories.

3.6 Traffic Scenarios

Following the independent variables, four experimental conditions can be determined, namely **A:LxN**, **B:LxY**, **C:HxN** and **D:HxY** for groups I and II, resulting in a total of eight conditions. Two repetitions of each condition were performed. In order to prevent recognition, dummy scenarios were used between actual measurement scenarios and measurement scenarios were rotated 180 degrees. In order to prevent confounds, only one type of conflict geometry was used for the actual measurement points, namely a converging track situation. The resolution advisories tried to solve conflicts in an efficient matter, whilst trying to steer the aircraft as much as possible toward their designated exit points.

4. RESULTS

4.1 Correct Accept/Reject Score

The correct accept and reject scores are an indication of the compliance to an advisory, implicitly indicating the trust the participants had in the SSD and the advisory. An initial analysis was performed by plotting the cumulative accept and reject scores, as shown in Figs. 4 and 5.

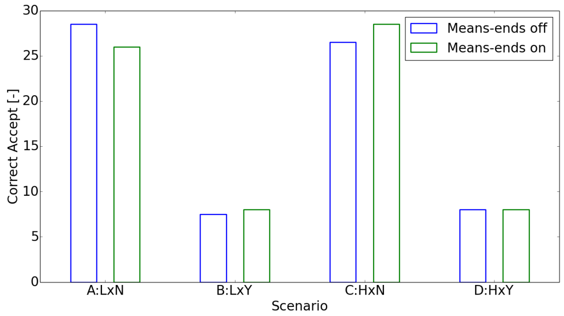


Fig. 4. Cumulative number of correct accept advisories

At first sight, a trend can be spotted between the presence and absence of a sensor failure. There does not seem to be a significant difference for the accept scores between groups, which is confirmed by a Kruskal-Wallis test. However, a Friedman test shows there are within-subjects interactions ($\chi^2(1) = 43.596, p < 0.001$). A post-hoc analysis using a Wilcoxon Signed Ranks test for pairwise comparison show there are significant effects between scenarios A-B ($Z = -3.549, p < 0.001$), A-D ($Z = -3.555, p < 0.001$), B-C ($Z = -3.542, p < 0.001$), C-D ($Z = -3.551, p < 0.001$), clearly indicating a negative effect of sensor failure on advisory acceptance.

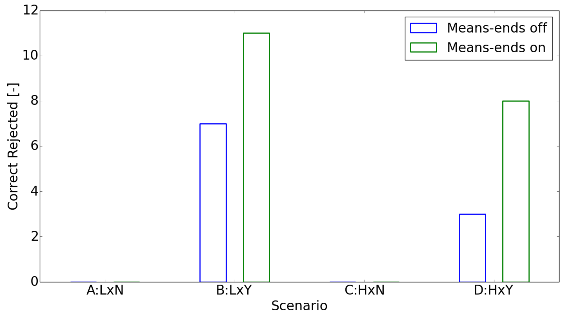


Fig. 5. Cumulative number of correct rejected advisories

For the reject scores, there seems to be some trends between groups, and between sensor failure conditions. However, a Kruskal-Wallis test showed there is no significant difference between the two group distributions for scenarios B ($\chi^2(1) = 1.049, p = 0.306$) and D ($\chi^2(2) = 4.000, p = 0.406$). A Friedman test showed within-subject interaction effects ($\chi^2(1) = 4.445, p = 0.035$), as confirmed by a post-hoc Wilcoxon test ($Z = -2.365, p = 0.018$). This means that there is a negative effect of scenario complexity on correctly rejected advisories.

4.2 Agreement Rating

Agreement ratings were given for every advisory to indicate the level of agreement with the given resolution, measuring yet another dimension of trust in the system. These scores are normalized and analyzed using boxplots, shown in Fig. 6. A trend can be seen between the presence or absence of a sensor failure, showing less agreement to an advisory in case of a sensor failure.

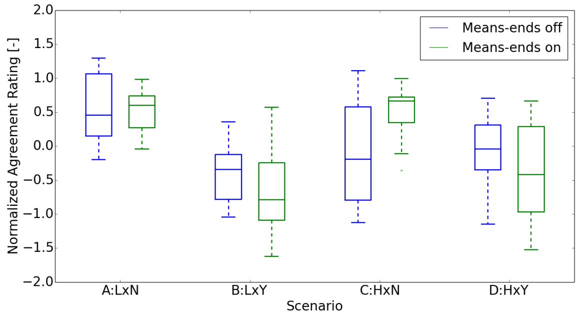


Fig. 6. Normalized agreement rating

Using a 2x2x2 three-way mixed Analysis of Variance (ANOVA) test a significant main effect of sensor failure ($F(1, 14) = 19.283, p = 0.001$) is found on agreement rating. However, neither main effects of group and complexity are found, nor any interaction effects.

4.3 Sensor Failure Detection

Sensor failure detection is mainly measured through verbal comments. Only when the participants found the correct aircraft subjected to sensor failure, this would be recorded. The cumulative numbers of detection are shown in Fig. 7. There seem to be an indication of different distributions, both between groups as between complexities.

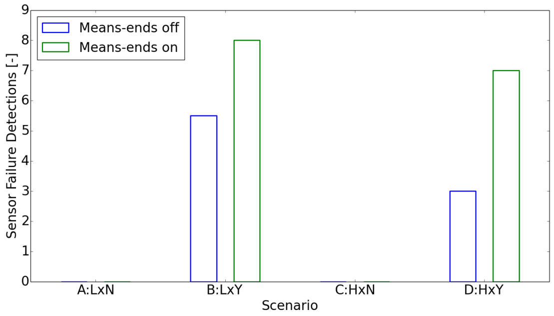


Fig. 7. Sensor failure detection

After performing a Kruskal-Wallis test it is shown that both conditions show a significant difference between the two groups, namely for scenario B:LxY $\chi^2(2) = 4.923, p = 0.027$ and for scenario D:HxY $\chi^2(2) = 5.208, p = 0.022$. It can be seen that this difference is relatively larger for the high complexity case. A Friedman test also indicates a significant difference within subjects, where a higher complexity would give a lower sensor failure detection rate $\chi^2(1) = 4.500, p = 0.034$.

4.4 Response Time

The advisory response time was recorded as the time between the start of the advisory and accepting or rejecting. The results are plotted in Fig. 8. At first sight it can be observed that group II has a consistently longer response time compared to group I, most likely caused by the elaborate usage of the means-ends functionality. No clear distinction between complexities or failure can be observed from the plots.

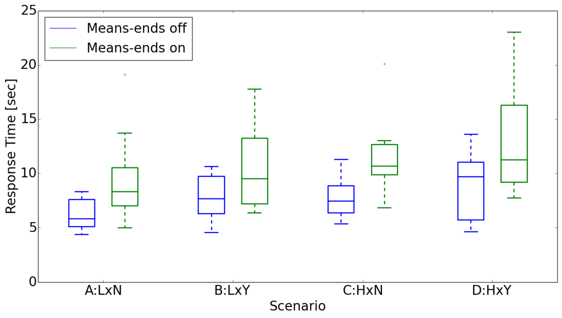


Fig. 8. Advisory response time

A 2x2x2 three-way mixed ANOVA test shows a hint of different distributions ($F(1, 14) = 206.249, p = 0.051$), between group I and II, however this is not significantly sufficient from a statistical point of view. A significant main effect of complexity ($F(1, 14) = 14.328, p = 0.002$) and sensor failure presence ($F(1, 14) = 10.233, p = 0.006$) is found. No interaction effects were found between group, complexity and sensor failure presence.

4.5 Observed Control Strategies

Several control strategies have been identified, of which the most frequently occurring ones are briefly addressed in this section.

(1) Check advisory with SSD, link conflict zones to the aircraft (mostly group II), and crosscheck this with radar image
(2) Reject correct advisories and manually redirect (sometimes temporarily into a conflict zone) for larger margins
(3) Accept/reject advisories such that an aircraft is manually controllable, then implement manual resolution
(4) Reject correct advisory, followed by a manual resolution which was approximately the same as the given advisory

Through observations it was seen that there were quite some different strategies used, which have influenced the results. Some participants had more experience with ATC and had knowledge about the rules of the air, which influenced their judgment on the advisories. In some cases these participants would prefer dual-aircraft resolutions, however only one aircraft per conflict was enabled for manual control after the advisory. It must be noted that these participants were balanced over the two groups.

Most participants also indicated to favour manual heading clearances over speed commands. The manual commands led in some cases to self-induced conflicts, resulting in a higher traffic complexity than initially designed. Due to these different control strategies, it is not always apparent to correlate a rejected advisory with the benefit of explicitly representing the means-ends links. That is, the advisory could be rejected because a sensor failure was detected, or because the advisory was non-conformal to the desired resolution of the participant, or due to general reluctance to use automation at all. However, observations and verbal comments helped identifying which participants correctly detected the aircraft subjected to sensor failure.

5. DISCUSSION

Looking at the results, there are signs that explicit representation of the means-ends relations enables improved sensor failure diagnosis, especially in more complex traffic scenarios. It is statistically proven that the group with means-ends on performed better at sensor failure detection than the group without explicit means-ends links. Verbal comments indicate that both groups could recognize something was incorrect during sensor failure, however only the means-ends group was better able to identify the aircraft with ADS-B failure. The additional information provided by explicitly visualizing the relations seems to help controllers in correctly identifying problems, and manually coming up with correct solutions.

Furthermore, it is shown that improved fault diagnosis is more pronounced in unstructured, complex situations. This is also partially supported by the correct reject scores, showing a trend between the groups, even though there is no confirmed statistical significant difference in distributions. A higher complexity (unstructured scenario) led to more overlapping conflict zones, which caused more trouble for group I.

The correct accept/reject scores showed no significant differences between the two participant groups. The high rate of (incorrect) advisory rejections can be explained by a number of reasons. First, there might be a lack of strategic conformance with the automation. Conformal resolution advisories are generally more easily accepted by the participant. For example, if there is a mismatch in resolution strategy between human and machine, the human tends to reject this advisory (Westin et al. (2013)). This conformance is personal and subjective, and might not always be the most optimal solution. The resolutions given in this experiment were in some cases due to geometrical restrictions not according to the rules of the air and thus non-conformal with some participants. Ideally, first a strategic conformance study should be performed among participants before the actual experiment, such that the preferred manual resolutions can be used to construct advisories, filtering out any conformance confounds (Westin et al. (2013)).

Another reason for the number of (incorrect) advisory rejections could be related to trust issues that arose when participants were subjected to a sensor failure (Lee and See (2004)). Some participants became quite hesitant to accept advisories based on the automation, even in scenarios without sensor failures. This happened specifically with those participants who had sensor failures in the first scenarios, indicating some order effects.

6. CONCLUSION

This paper presented the investigation of the effects of visualizing means-ends relations (between aircraft and their conflict zones) in an ecological interface on sensor failure detection. This was done by conducting a human-in-the-loop experiment using a novel Air Traffic Control (ATC) ecological interface for Conflict Detection & Resolution (CD&R) called the Solution Space Diagram (SSD). It was hypothesized that enabling the visualization of these links helps controllers detect and identify sensor failures, improving their performance. Results reveal that there are strong indications that the added functionality of means-ends relations positively contributes to sensor failure detection and diagnosis. In some cases there were incorrect rejected advisories due to a lack of advisory conformance with the participants, or general reluctance to accept automation. However, the means-ends relations are extensively used in order to inspect the SSD and their advisories, resulting in a longer advisory response time and more accurate sensor failure detection performance. Even though this research is an important first step, further research is required to answer remaining questions and investigate means to address trust and conformance issues.

REFERENCES

Ali, B., Majumdar, A., Ochieng, W.Y., and Schuster, W. (2013). ADS-B: The Case for London Terminal Manoeuvring Area (LTMA). In *Tenth USA/Europe Air Traffic Management Research and Development Seminar*, 1–10.

Borst, C., Westin, C., and Hilburn, B. (2012). An Investigation into the Use of Novel Conflict Detection and Resolution Automation in Air Traffic Management. *SESAR Innovation Days*, (November).

Borst, C., Flach, J.M., and Ellerbroek, J. (2015). Beyond Ecological Interface Design: Lessons From Concerns and Misconceptions. *IEEE Transactions on Human-Machine Systems*, 45(2), 164–175. doi: 10.1109/THMS.2014.2364984.

Burns, C.M. (2000). Putting it all together: Improving display integration in ecological displays. *Human Factors*, 42(2), 226–241.

Cedrini, V., Zacchei, M., Zampognaro, V., Ente, E., Assistenza, N., Spa, V., and Agri, V. (2010). ADS-B 1090ES Implementation : the CRISTAL-MED Project. 3, 1–5.

ICAO (2013). The Ninth Meeting of the Southeast Asia and Bay of Bengal Sub-regional ADS-B Implementation Working Group (SEA/BOB ADS-B WG/9). Technical report, Beijing.

Lee, J.D. and See, K.A. (2004). Trust in automation: designing for appropriate reliance. *Human factors*, 46(1), 50–80.

Parasuraman, R., Sheridan, T.B., and Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and humans*, 30(3), 286–97.

Rekkas, C. and Rees, M. (2008). Towards ADS-B Implementation in Europe. *Tyrrhenian International Workshop on Digital Communications - Enhanced Surveillance of Aircraft and Vehicles*, 1–4.

St-Cyr, O., Jamieson, G., and Vicente, K. (2013). Ecological Interface Design and Sensor Noise. *International Journal of Human-Computer Studies*, 71(11), 1056–1068. doi:10.1016/j.ijhcs.2013.08.005.

Westin, C.A.L., Hilburn, B.H., and Borst, C. (2013). The Effect of Strategic Conformance on Acceptance of Automated Advice: Concluding the MUFASA Project. *Proceedings of the SESAR Innovation Days*, 3(November).

Zhang, J., Liu, W., and Zhu, Y. (2011). Study of ADS-B Data Evaluation. *Chinese Journal of Aeronatics*, 24(4), 461–466.