EVALUATION OF AN ALTERNATE AIRPORT ASSISTANT PROTOTYPE BY MEANS OF AN ONLINE PILOT STUDY

A. Paul*, J. I. González Cabeza*, G. Schmitz*
* German Aerospace Centre, Institute of Flight Systems, 38108 Brunswick, Germany

Abstract

The project Next generation Intelligent Cockpit (NICo) of the German Aerospace Centre (DLR) aimed at developing assistance systems in the context of single pilot operations (SiPO). Failure management is one challenge of SiPO because it places the highest demands on the crew. In order to make an informed decision, the effects of a failure must be determined quickly. Automated provision of relevant facts in the respective situation can relieve the pilots of bothersome tasks like gathering data and combining information. In previous studies DLR investigated criteria pilots apply when considering options for alternate airports in case of diversion from the planned destination. Criteria including distance to airport, fuel on arrival, stop margin and crosswind were identified as relevant for pilots deciding where to divert. Based on the findings, DLR designed a prototypical graphical user interface (GUI) that should assist pilots with finding suitable airports. It was tested during a pilot study carried out remotely in scope of a master's thesis.
The aims of this study were to assess the feasibility of the GUI's various features by gaining pilot feedback on its usability and improvement possibilities. 14 pilots with current or past Airbus A320 type rating performed two online exercises and completed associated questionnaires. The pilots made use of different GUI features when considering their options to divert after encountering malfunctions that were not time-critical. In an analysis of results, it became evident that pilots rated the prototype overall favourably, although they approved of the functionalities to varying degree. The feedback is valuable for the ongoing development of the application and encourages to consider the duration of the mission, i.e. short-range or long-range flights.

Keywords
Decision-making; single pilot operations; assistance system; FORDEC

NOMENCLATURE

Abbreviations

AI	Artificial Intelligence
ASAP	As Soon As Possible
ATPL	Airline Transport Pilot Licence
CAT	Category
CB	Cumulonimbus
DLR	Deutsches Zentrum für Luft- und Raumfahrt, German Aerospace Centre
EASA	European Aviation Safety Agency
EFOB	Estimated Fuel On-board
eMCO	Extended Minimum Crew Operations
FORDEC	Facts, Options, Risks & benefits, Decision, Execution, Check
GUI	Graphical User Interface
ICAO	International Civil Aviation Organisation
ILS	Instrument Landing System
MCDM	Multi-Criteria Decision-Making
METAR	METeorological Aerodrome Report
MPO	Multi-Pilot Operations
NICo	Next Generation Intelligent Cockpit
QCA	Quantitative Content Analysis
RCO	Reduced Crew Operations
SiPO	Single Pilot Operations
TAF	Terminal Aerodrome Forecast
TCU	Towering Cumulus
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
VCP	Virtual Co-Pilot

doi: 10.25967/650409

1. INTRODUCTION

As global air traffic is forecast to grow significantly in the future and qualified aircrew is becoming scarce [1], concepts of reduced crew operations (RCO) have been investigated [2]. In recent times, the industry has been focussing on developing ideas to minimise the number of active crew members during cruise while one crew member is resting (extended Minimum Crew Operations, eMCO) [3]. Further reductions would lead to single pilot operations (SiPO) albeit neither concept is yet certified for EASA CS-25 aircraft. The project Next generation Intelligent Cockpit (NICo) of the German Aerospace Centre (DLR) aimed at developing assistance systems in the context of SiPO to support and relieve pilots in phases of high workload [4].

Such a phase of high workload can be encountered when unforeseeable incidents (e.g. medical emergency, system failure, weather deterioration, airport closure) occur during a flight. If these require a landing at an alternate airport that is not included in the flight plan, the pilots need appropriate systems to identify a suitable alternate airport. The variety of possible incidents, the surrounding topography, and the prevailing weather conditions make it impossible to foresee every possible scenario during flight preparation. There is therefore a need to be able to react to incidents during the flight and to eventually identify an alternate airport.

According to EASA Air Operations [5], a distinction is made between four categories of alternate airports. Firstly, so-called destination alternates are taken into consideration if a landing at the destination airport is not possible. En-route alternates are located along the flight route and are considered, for example, in the event of technical defects or incidents in the cabin during the flight. Reduced contingency fuel alternates represent the third category and are taken into consideration when at a pre-determined point on the flight route there were not enough fuel available to land at the actual destination airport. Fourthly, the so-called take-off alternate airports are approached if a malfunction occurs during take-off. For alternate airports which are included in the flight plan, relevant airport and weather data is obtained in preparation for the flight and discussed in the briefing. During the flight, weather conditions and other relevant data are constantly monitored for these alternate airports [6].

Pilots are trained to adhere to decision-making schemes, e.g. FORDEC. The mnemonic acronym FORDEC abbreviates the steps of collecting Facts, generating Options, analysing Risks and benefits, making a Decision, Executing it and Checking its adequacy over time [7]. The NICo project introduced the concept of a Virtual Co-Pilot (VCP) to explore how automation can assist flight crews or single pilots in diagnosing failures and making decisions. According to recent expert interviews, pilots would welcome additional support, particularly with the FORDEC decision-making steps [8]. Notwithstanding, the experts emphasised that the final decision and its execution must remain the responsibility of the human crew or pilot, not an automated system. The primary goal for developing the VCP is to enhance safety in both multicrew and single-pilot flight operations. This ensures that all VCP functions can be applied broadly, regardless of whether single pilot operations are deemed feasible in the future.

A literature review revealed several decision support tools that propose and rank alternate airports. They use different underlying algorithms, and vary by the criteria considered as well as the user interface.
Atkins *et al.* [9] suggested an algorithm that uses eight parameters in a weighted sum criterion, including runway dimensions, distance, and wind velocities. Grzybowski and Szpakowska-Peas [10] combined multi-criteria decision-making (MCDM) methods with membership functions that map the value of each criterion to a weighting. Coombes *et al.* [11] suggested multi-criteria decision-making Bayesian Networks to handle uncertainty in the criteria employed. A dynamic re-routing tool was created using deep learning to find the best alternative flight paths based on past flight data and current weather [12]. This tool ranks alternatives, highlighting advantages with colour codes. However, pilots in feedback sessions asked for more insight into how the tool makes its suggestions to increase their confidence in using it.

Zhang *et al.* [13] developed prototypes to explore how AI can be used in decision support tools for pilots. These tools not only suggest and rank alternate airports based on predefined criteria but also monitor data to alert pilots to important information. The study found that pilots are unlikely to trust a decision support tool easily, and they also resist attempts to overtly calibrate their trust during the decision-making process. Instead, the researchers recommend focussing on supporting pilots before and during the decision-making process to help them maintain situational awareness, which is critical for making good decisions in complex and dynamic situations. They also emphasise that decision-making is a more nuanced process than just choosing between options, and more research is needed to understand real-world decision-making environments.

Mainly to assist with Facts, Options, and Risks categories of FORDEC when exploring alternates for a diversion, DLR has created a GUI prototype. Suitability of airports was determined by four key factors: distance, fuel on arrival, stop margin, and crosswind. Initially, these factors were given equal weight using a multi-criteria decision-making approach. Findings from a previous pilot survey confirmed the importance of these criteria in determining where to divert [14]. The aims of the present study were thus to assess the feasibility of the prototype and to get pilot feedback on its usability.

Concerning the pilots the following research hypotheses were proposed:

1) Pilots will rate the application differently according to their age.
2) Pilots will rate the application differently according to their flying experience.
3) Pilots will rate the application differently according to their rank (i.e., first officer or captain).
4) Pilots will rate the application differently according to their operational range (i.e., short-haul or long-haul aircraft).
5) Pilots will rate the different functions useful to varying degree.

A difference in assessment between the age groups is assumed, as it can be argued that the group aged under 55 is more prone to use electronic devices in their private lives than the older participants. It is also possible to categorise pilots according to the type of flight route. It should be mentioned here that long-haul pilots typically worked in short- and medium-haul operations before being type-rated on long-haul aircraft. With regard to the evaluation of the GUI prototype, a difference is assumed between the different flight route types, as long-haul pilots in particular have to deal more with the issue of alternate airports during flight preparation than short-haul pilots. As long-haul flights often fly over large expanses of water or uninhabited areas, suitable alternate airports are sometimes far away from the planned flight route. In addition, the long-haul aircraft place special requirements on airports as they require a longer runway for take-off and landing. Furthermore, the airport requires the appropriate equipment to be able to handle a long-haul aircraft. The GUI prototype itself comprises different functions that may prove more or less helpful in decision-making. A difference between functions is assumed with regard to support in decision-making, as the functions cover different aspects. While some functions map a filtering of the data, other functions prepare the data graphically. It is of interest whether these functions will be rated differently.

2. METHODS

2.1. Participants

A total of 14 pilots from various airlines aged between 25 and 70 took part in the study. The participants were all male and worked in an MPO cockpit configuration. Of the 14 pilots, 13 completed the preliminary questionnaire and the final questionnaire. Still, it was not the same pilot who did not answer both questionnaires. The average age of the participants can

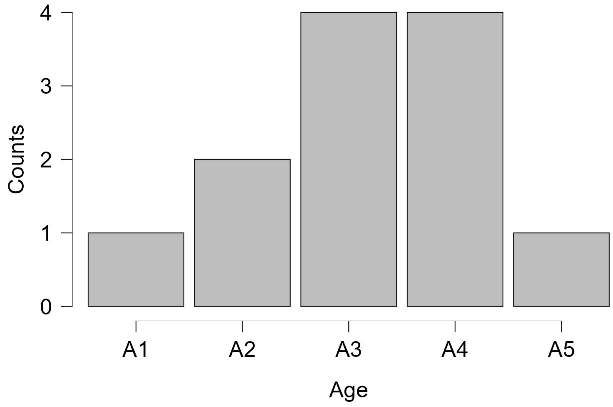
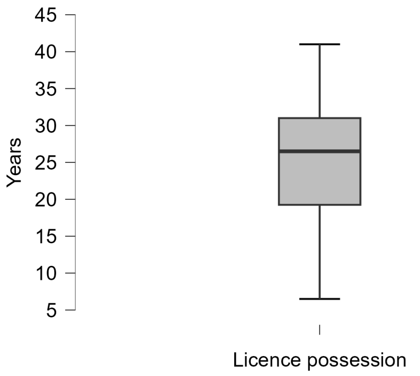


FIG 1. Participant age distribution; A1 (25-34 y.), A2 (35-44 y.), A3 (45-54 y.), A4 (55-64 y.), A5 (>65 y.)

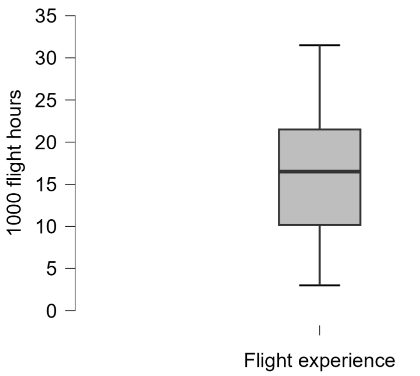
only be determined approximately, as the preliminary questionnaire did not ask for a specific number, but rather for an age group. The average age is calculated as 48.7 years, taking into account the mean value of each age group. The distribution of the age groups can be seen in Figure 1.

Participants were recruited by e-mail as they had expressed their interest in previous studies and had a valid Airbus A320 type rating at the time of inclusion in the mailing list. However, as this entry was possibly made several years ago, it cannot be concluded that the participants held a valid type rating for the A320 at the time of this study. The fact that the participants also included pilots who stated that they were predominantly involved in long-haul operations supports this statement. Specifically, six of the 13 pilots were operating long-haul flights (>2000 nm), while five pilots flew short-haul (<800 nm). Two of the 13 pilots stated that they were flying mainly medium-haul (800-2000 nm). Even though some pilots did not have a valid type rating for the A320 at the time of the study, it can be assumed that every participant in the study was familiar with the operation of the A320 due to prior experience. With the exception of one pilot, all study participants were mainly active in commercial aviation. The remaining one operated mainly in business aviation. Twelve of the 13 pilots stated that they work in passenger air transport, while one worked in both cargo and passenger air transport.

The participants have held their ATPL for an average of 25 years and have completed an average of 15,600 flight hours during this time. The period of possession of the ATPL licence varied from 6.5 to 41 years, with the majority of pilots holding the licence for 19.25 to 31 years. The number of flight hours completed ranged from 3000 to 31500. Most of the pilots had accumulated between 10000 and 21500 flying hours. Their distribution can be seen in the boxplots in Figures 2a and 2b. The participants included ten captains and three senior first officers. The latter is a rank that authorises the first officer to take over the captain's re-



(a) Box plot of licence possession time



(b) Box plot of flight experience

FIG 2. Pilot's experience distributions

sponsibilities during the cruise flight phase which is particularly important for long-haul flights.

Nine of the 13 pilots stated that they had additional qualifications including activities as flight instructors, instrument rating instructors, synthetic flight instructors, or type rating instructors, as well as activities as test pilots, performance engineers, or consultants for human factors in aviation. Five of the 13 pilots said that they had already taken part in DLR studies.

2.2. Materials

The apparatuses employed in the research included the graphical user interface prototype and a videoconferencing software as well as two standard Windows computers with mouse, keyboard, and headphones. Video and audio was recorded after the participants gave their explicit consent. Questionnaires were used to collect socio-demographic data and information on professional experience or any additional qualifications. In order to obtain statistical data on the prototype, the participants answered a questionnaire upon concluding the interview. It included an evaluation of the individual functions of the prototype with regard to the support and user-friendliness of the respective function as well as general questions on selected functions. The participants' statements were recorded on Likert scales with an even number of answer options. They offer the advantage of capturing a tendency by removing the neutral central option [15].

The GUI prototype shown in Figure 3 was implemented in MATLAB R2007b. The calculations are based on handbook methods for the Airbus A320 aircraft type, which was chosen due to its commonality and also prior experience at DLR (part of the DLR research fleet). Geographically, the prototype is limited to the European continent at the current stage of development. The user can select a date in the past on which the planned flight should have taken place. In addition, the system failure to be considered can be selected. The stored database covers technical malfunctions in the flap system, the hydraulic system, and the flight control system. Taking into account the consequences of a malfunction on aircraft performance is an elementary feature of the prototype. The aim of future development is to provide these data automatically by linking the GUI to the aircraft's systems.

The filter criteria implemented in the prototype are the result of the aforementioned preliminary study [14]. The user can filter by approach procedures, wind conditions and four operational criteria: stop margin, estimated fuel on arrival, crosswind per runway and great circle distance. A histogram can be generated for each criterion, displaying the number of runways that fulfil it. Histograms for three criteria can be seen in Figure 4. The representations on the map can be customised using the various map overlays. By default, the map displays the airports that fulfil the filter criteria selected by the pilot. Furthermore, the map displays a red aircraft symbol that indicates the position of the aircraft.

A ranking of airports is based on the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) which is a decision-making methodology that takes several criteria into account. This methodology approaches the ideal solution via the distance to the worst and best ideal solution [16]. It is important to note that each of the five factors (stop margin, EFOB on arrival, distance, crosswind, CB/TCU forecast) is equally weighted.

The top right field contains the ICAO code and the name of the airport as well as information about the flight and the time of arrival. Data on the weather at the airport at the time of arrival is displayed in the form of the current TAF. For Runway condition codes 1 to 6, the prototype calculates a stop margin using handbook methods that take into account the prevalent system failure. The available runways and associated approach procedures are highlighted with colours. Green runways indicate that these runways have a positive stop margin and are suitable for landing, taking into account the weather conditions. Runways coloured red are not suitable for landing according to various criteria. Reasons for a red

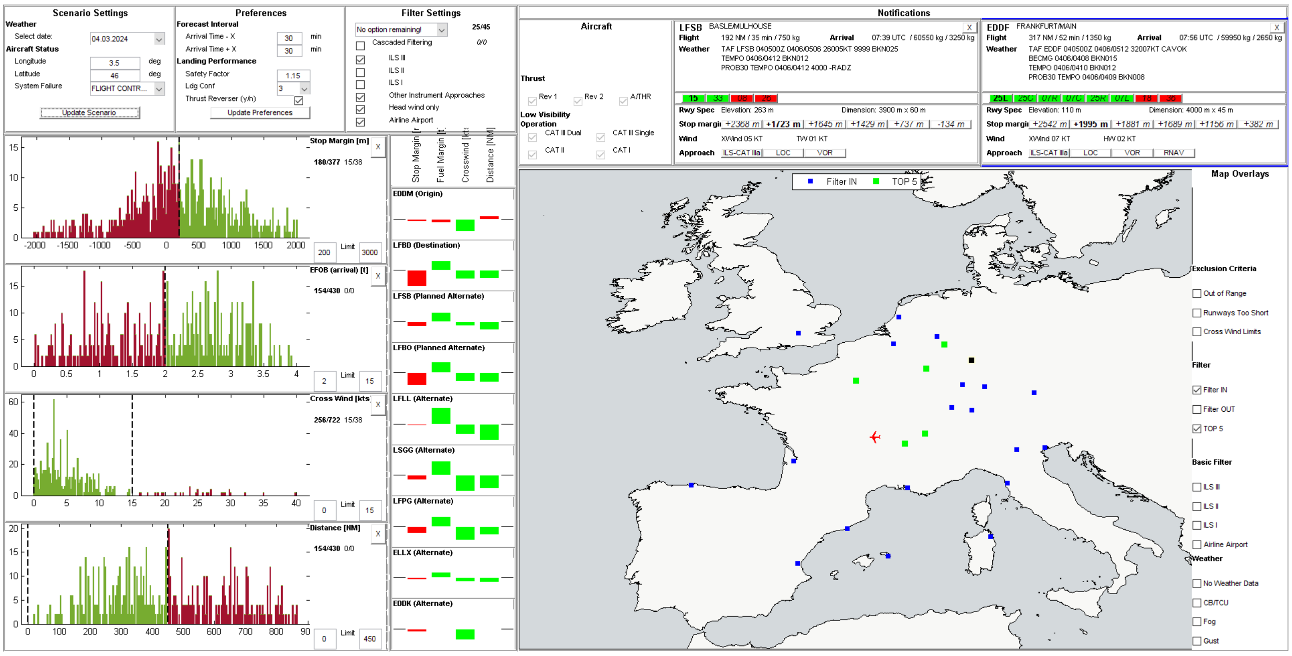


FIG 3. Overview of the GUI prototype

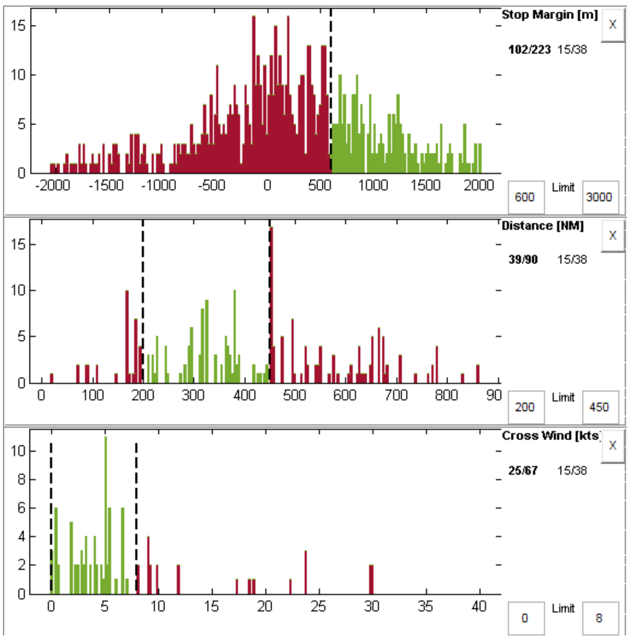


FIG 4. Histograms for Stop Margin, Distance and Cross-wind

indication can be a negative stop margin, a crosswind exceeding the limits, or weather conditions that are below the minima defined for the approach procedure. The cause for the colour coding is displayed upon hovering over the runway designator.

In the comparison view, as shown in Figure 5, the pilot can make a visual comparison of airports using bar diagrams. In the list there are the origin, destination and planned alternate airports as well as the top ranked airports according to the TOPSIS method. The comparison airport can be freely chosen and is referenced in the top right info box above the map. The direction of the bars and their colour are intended to assist the pilot in assessing the airports. For the stop margin and EFOB criteria, positive deflections are coloured green, while negative deflections are coloured red. Higher stop margins and fuel reserves on arrival are therefore rated positively, while a reduction in these is rated negatively. For the crosswind and distance criteria, this is the other way around. Negative deflections are coloured green, while positive deflections are coloured red. This is due to an increase in crosswind or distance to the alternate airport is rated negatively. For each column the bar height is relative to the largest value of the respective criterion, i.e. the indications are normalised.

2.3. Procedure

This subsection describes the rationale behind the decision to conduct the study as an online meeting rather than a face-to-face meeting. In terms of methodology, the face-to-face study offers the advantage that direct communication with the participants is possible, particularly transporting non-verbal information [17].

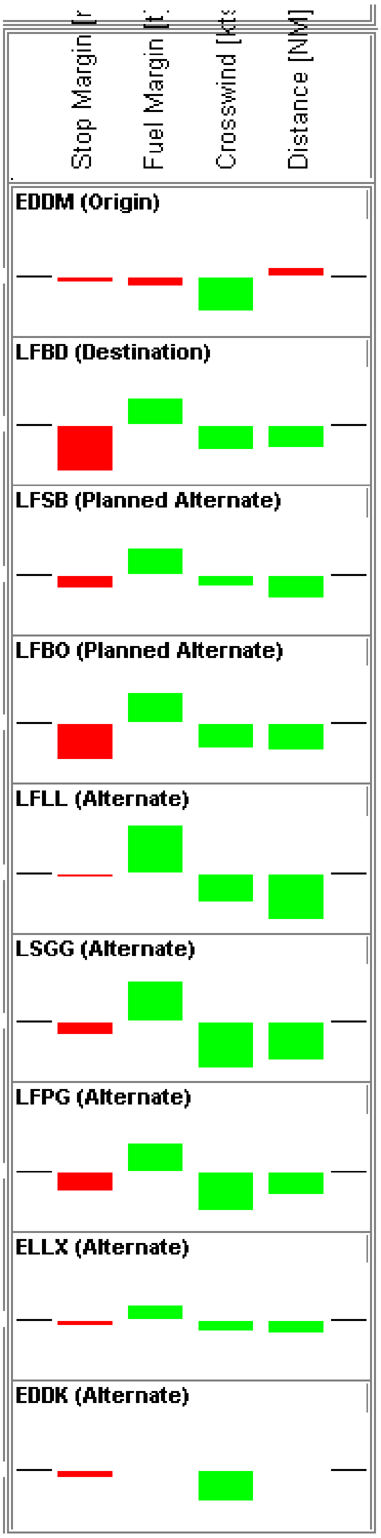


FIG 5. Comparison view, reference EDDF airport

Moreover, face-to-face studies are more attractive to participants as they are not as strenuous as online meetings. According to Buch and Schmitz [17], the loss of non-verbal communication in online meetings can also lead to silent breaks, which have no clear explanation. For a face-to-face study, a suitable location must be chosen in order to minimise the effort for the participants. The number of participants could be lower under these circumstances. This leaves an online meeting to conduct the study which offers the advantage of a high degree of flexibility for participants. They are more free to decide where and when to take part in the study. Still, conducting the study as an online meeting also has disadvantages. An online meeting can capture non-verbal communication only to a limited extent. Furthermore, direct interaction with the prototype cannot be realised. This is due to the fact that the participants do not have access to the required software.

With regard to the number of participants, a further aspect must be discussed. Usually, today a commercial aircraft is flown in multi-pilot operation configuration. This means that the processes in the cockpit are designed so that two pilots are responsible for the flight. According to Pohl [18], this has the advantage of reducing the phenomenon of confirmation bias (decision-making based on one's own expectations) and target fixation (fixation on a target without recognising other options). However, as conducting the study in MPO mode would tie up two pilots per experiment, the number of experiments would be halved.

In addition, the study focusses exclusively on the aspect of finding a suitable alternate airport, which is why the usual main tasks of the flight (aviate, navigate, communicate) do not play a role. Although the decision for an alternate airport is always made by two pilots in order to prevent the phenomena described above, each pilot is in general able to search for an alternate airport on their own. In addition, the realisation of two-pilot studies would have a negative impact on study planning. Eventually, it was also the long-term goal of the NICo project to investigate operation in single-pilot operation configuration, which additionally supports the realisation of a single-pilot study.

On the scheduled date, the pilots dialled into the web conference and, after a brief welcome, were introduced to the study. Two scenarios were employed in the research and were administered subsequently during the video conference.

A non-time-critical malfunction is assumed for the first scenario. According to the available aircraft documentation a jammed rudder malfunction was chosen. This is suitable for analysing the prototype, as it is a malfunction that limits the maximum permitted crosswind during landing to 15 knots due to the missing control authority during the de-crab manoeuvre and subsequent ground control. The restriction is to be interpreted as decisive for the selection of the alternate airport. In addition, only the rudder is affected by the system fault, which means that there are no further restrictions for the other aircraft systems.

Possible flight routes were limited to connections on the European continent. In order to give the participants sufficient time to search for an alternate airport, the fault should occur at a sufficient distance from the destination airport. Moreover, the flight must already be at an advanced stage in order to restrict a return to the departure airport. In light of these points, a flight from Stockholm (ICAO-code: ESSA) to Rome (LIRF) was assumed for scenario one. The rudder jam occurs when flying over the eastern German Baltic coast at cruising altitude. In order to analyse the use of the prototype, a day with high ground wind speeds over Germany was selected. The weather data were obtained for 21 December 2023 using the website Ogimet [19], where it is possible to display past weather situations and weather forecast. They are available in the form of the METAR (Meteorological Aerodrome Report) and TAF (Terminal Aerodrome Forecast), for each airport stored in the database.

The aim of the scenario is to analyse the prototype in critical situations. However, the chosen malfunction should not cause a time-critical situation as these are mostly associated with a "Land ASAP" warning which mandates for an immediate landing which results in the pilots' choosing the next suitable airport for landing without extensively using the decision support prototype. The scenario malfunction should therefore trigger a need to land, but this not instantaneously. For the second scenario a hydraulic fluid loss of one of the three A320's hydraulic systems was chosen. The aircraft remains manoeuvrable using the remaining two hydraulic systems, but the malfunction causes the following restrictions which cannot be compensated for using the remaining hydraulic circuits:

- Failure of the nose wheel steering
- Failure of the thrust reverser of engine 1
- Failure of the automatic brakes
- Failure of the mechanism for retracting and extending the landing gear
- Failure of the ability to land under CAT III Dual conditions

In particular, the failure of the hydraulic landing gear control affects the selection of the alternate airport, as the landing gear can be extended but cannot be retracted afterwards. If the pilots decided to abort the landing, the landing gear could not be retracted, which would significantly increase the aircraft's fuel consumption. This is relevant for the considerations in this scenario, as the pilots would thus be clearly committed to a region when choosing an alternate airport. A flight from Marseille (LFML) to Stavanger (ENZV) is modelled for the second scenario. This

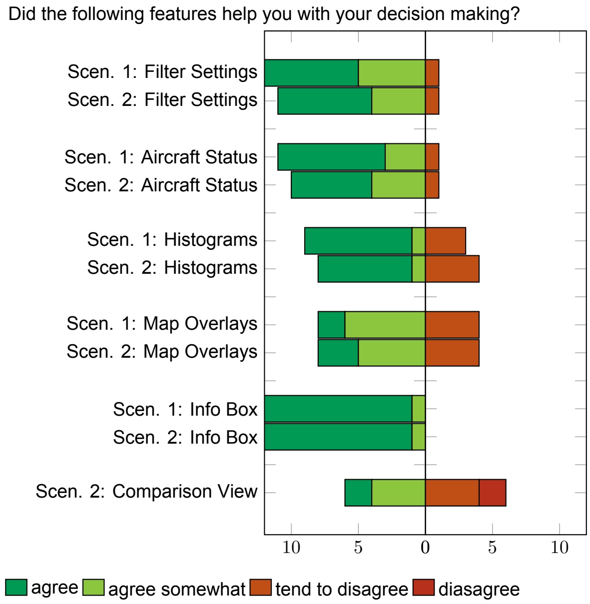


FIG 6. Rating of GUI features' usefulness

routing has the advantage that part of the flight runs over the North Sea leaving less options to divert than over land. Furthermore, the aircraft has already completed a large part of the flight when flying over water, meaning that most of the available fuel has been used up.

2.4. Measures

The statements in the questionnaire were answered using Likert scales with the following items: agree, agree somewhat, tend to disagree, do not agree. All participant entries were recorded anonymously. In addition to the questionnaires, the qualitative content analysis (QCA) method according to Mayring was used to analyse the study. This is a method for analysing communication material or texts [20]. QCA, which is particularly suitable for analysing the content of data, can be used to summarise the participants' statements relevant to the respective function. For this purpose, a transcript with corresponding time references was produced using the recording after the study. The transcribed statements were categorised according to Mayring's content analysis. In this case each function of the prototype represents a category. The statements on the categories are summarised for all participants and discussed as part of the analysis.

3. RESULTS

3.1. Differences between groups

To test for differences between pilot age, rank, operational type, and flight experience regarding

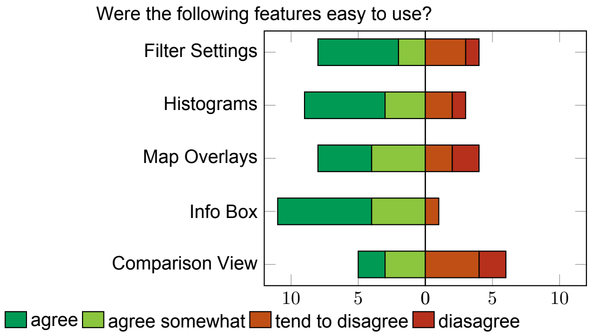


FIG 7. Rating of GUI features' user-friendliness

the responses, a series of Kruskal-Wallis tests was employed. There was no significant difference concerning the ranks of the items on the Likert scales.
A series of logistic ordinal regression analyses was performed to investigate the relationships between ratings of the prototype and pilot age or flight experience, respectively. No significant correlations were supported by the results.

3.2. Response frequencies

Figures 6 and 7 show the frequencies of responses to some questions contained in the questionnaire that was administered upon concluding the experiment. It can be seen that most features were rated positively, although the comparison view was found less helpful. Regarding the usability, the comparison view was rated not as user-friendly as the other features.

3.3. Qualitative content analysis (QCA)

The pilots commented on various aspects relating to the selectable filters and filtering in general. Ten of the 14 pilots noted a need for improvement in the selection of the available approach procedure. 13 of the 14 pilots also noted that the filter option of the airline airport does not sufficiently cover the operational airline needs. Proposed additional filter criteria were availability of airline-specific maintenance, medical care options, availability of aircraft ground handling for the aircraft type, and availability of multiple independent runways as a backup if one should be unexpectedly closed on short notice.

Ten of the 14 pilots stated that highlighting the departure and destination airports in colour would be a useful addition in order to be able to identify them more quickly on the map. Some pilots also stated that this highlighting should also be applied to the airline's hub airports. Moreover, twelve of the 14 pilots stated that they thought it would be useful to integrate the planned flight route of the flight into the map so that the displayed alternate airports can be better viewed in relation to the current flight route. Moreover, ten of the 14 pilots stated that labelling would be necessary to identify the airports on the map, which are only represented by a box. When asked whether a search function that allows airports to be identified and called up by entering the location or the IATA code would be a useful addition, the responses varied. Five pilots were in favour of implementing a search function, while three pilots explicitly rejected it.

With regard to overlays, five of the 14 pilots were in favour of taking the topography into account and integrating it into the map, which could also be implemented as an overlay. Furthermore, four pilots mentioned that a distance display from the current position of the aircraft in the form of circles would be a useful extension.

Concerning the comparison view, which was only used in the second scenario, the qualitative content analysis of the statements from the online meeting revealed various aspects for expanding this function. Four of the 14 pilots stated that the normalisation makes the interpretation of the bars more difficult. One pilot stated that excluding airports from the comparison view would be a useful enhancement. This would allow airports that should no longer be considered for a landing to be excluded and the normalisation of the bars to be updated. The majority of pilots reported that the comparison view is very complex and that the operation and interpretation of the results require a certain amount of routine.

4. DISCUSSION

4.1. Analysis

None of the research hypotheses could be supported or disproved statistically. Due to the limited sample size, also the number of individual responses was low. Statistically, there was no significant difference in responses for the pilots grouped by age, experience, operational range, or rank.
Pilots rated most features of the prototype favourably, albeit they stated that they were not always very intuitively to interact with. The response frequencies analysis shows that there was consensus among the participants for some questions, e.g. "How do you rate the GUI's support in choosing an alternate with regard to the second scenario?", that was answered favourably by eleven pilots. However, pilots were divided when answering other questions, e.g. when rating the user-friendliness of the prototype's features, cf. Figure 7.

4.2. Limitations and future research

With regard to the study participants, it should be noted that five of the 14 participants were over 55 years of age. Notwithstanding, it can be assumed that some of these pilots were no longer actively flying at the time of the study as the mandated retirement age for pilots, for instance in the USA is 60 years [21]. This circumstance is relevant because these pilots may have taken the study as an opportunity to find out about the current state of research and were less interested in the actual system. Nevertheless, no conclusions can be drawn from this regarding the quality of the statements made by these pilots. From this point of view, it would also have been interesting to ask in the questionnaire whether the pilots were still in active flying service and whether there was a measurable difference between the statements of the active and inactive pilots. In addition, the statements and answer options recorded in the questionnaire must be critically scrutinised. It should be noted that the use of Likert scales with four possible answers has advantages and disadvantages. The items used in the study were two positive and two negative answer options and the pilots also had the possibility to abstain from answering by choosing ”no answer”. Still, this makes it difficult to identify extremes in the pilots' statements. The use of a finer gradation of response options may also be of interest.

In light of the methodology used, it should be noted that there are various other options for evaluating the use of the prototype. One example is eye tracking, which can be used to record eye movements and gaze directions. Measuring the time taken to process the scenarios could also be used to identify particularly helpful functions. It should also be mentioned that, due to the limitations of the prototype, the scenarios represented two short-haul flights, which do not correspond to the usual flight route of long-haul pilots. Nevertheless, by selecting the flight over the North Sea, an attempt was made to simulate the overflight of large expanses of water, which tends to correspond more closely to the typical flight route of long-haul pilots. It should also be emphasised that the scenarios included two non-time-critical malfunctions because of which this study does not allow any conclusions to be drawn about the use of the prototype in time-critical situations.

Future research opportunities include the investigation of the prototype with regard to eMCO-SiPO concept currently being evaluated by EASA [2] . It could be examined the extent to which the prototype is able to improve the situational awareness of a pilot returning to the flight deck from a rest period. The integration of the prototype in a realistic flight deck (e.g. simulator) can also be of interest, such that basic flying tasks (aviate, navigate, communicate) have to be executed while interacting with it. One could further investigate the use of the prototype in a multi-crew flight deck where both pilots utilise it.

4.3. Conclusion

14 pilots with ATPL licence performed an exercise to assess the prototype of an application in a study carried out remotely. The participants rated the prototype overall favourably, however, no significant differences in ratings could be found with respect to pilot age, experience, or rank.
Quantitative content analysis revealed valuable pilot feedback concerning various aspects of the prototype. Based on this the application could be improved to better address the requirements in real-life operations.

Contact address:

joan.gonzalezcabeza@dlr.de

References

[1] D. Hulst, *Boeing Commercial Market Outlook*, Boeing, Ed., 2024. [Online]. Available: https://www.boeing.com/content/dam/boeing/boeingdotcom/market/assets/downloads/2024-cmo-executive-presentation.pdf.

[2] R. E. Bailey, L. J. Kramer, K. D. Kennedy, C. L. Stephens, and T. J. Etherington, “An assessment of reduced crew and single pilot operations in commercial transport aircraft operations,” in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, 2017, pp. 1–15. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8101988 (visited on 08/21/2025).

[3] D. Ziakkas, A. Plioutsias, and K. Pechlivanis, “Artificial Intelligence in aviation decision making process.The transition from extended Minimum Crew Operations to Single Pilot Operations (SiPO),” in *Artificial Intelligence and Social Computing*, ser. AHFE International, AHFE International, 2022. DOI: 10.54941/ahfe1001452.

[4] D. Niedermeier, A. Papenfuß, and M. Wies, “Simulator Study on a Spatially Separated Airline Crew in a Complex Decision-Making Scenario,” in *AIAA AVIATION 2023 Forum*, Reston, Virginia: American Institute of Aeronautics and Astronautics, 2023, ISBN: 978-1-62410-704-7. DOI: 10.2514/6.2023-3474.

[5] European Aviation Safety Agency, *Easy Access Rules for Air Operations*, 2025. [Online]. Available: https://www.easa.europa.eu/en/document-library/easy-access-rules/easy-access-rules-air-operations-regulation-eu-no-9652012.

[6] D. O'Hare and N. Stenhouse, “Under the weather: an evaluation of different modes of presenting meteorological information for

pilots," *Applied ergonomics*, vol. 40, no. 4, pp. 688–693, 2009. DOI: 10.1016/j.apergo.2008.06.007.

[7] H. Soll, S. Proske, G. Hofinger, and G. Steinhardt, "Decision-Making Tools for Aeronautical Teams: FOR-DEC and Beyond," *Aviation Psychology and Applied Human Factors*, vol. 6, no. 2, pp. 101–112, 2016, ISSN: 2192-0923. DOI: 10.1027/2192-0923/a000099.

[8] F. Zinn, J. della Guardia, and F. Albers, "Pilot's Perspective on Single Pilot Operation: Challenges or Showstoppers," *Lecture Notes in Computer Science*, vol. 14018, pp. 216–232, 2023, ISSN: 0302-9743. [Online]. Available: https://elib.dlr.de/194502/.

[9] E. M. Atkins, I. A. Portillo, and M. J. Strube, "Emergency Flight Planning Applied to Total Loss of Thrust," *Journal of Aircraft*, vol. 43, no. 4, pp. 1205–1216, 2006, ISSN: 0021-8669. DOI: 10.2514/1.18816.

[10] P. Grzybowski and E. Szpakowska-Peas, "Flight reconfiguration system – an emergency system of the future," *Aircraft Engineering and Aerospace Technology*, vol. 92, no. 9, pp. 1393–1399, 2020, ISSN: 1748-8842. DOI: 10.1108/AEAT-03-2020-0052.

[11] M. Coombes, W.-H. Chen, and P. Render, "Site Selection During Unmanned Aerial System Forced Landings Using Decision-Making Bayesian Networks," *Journal of Aerospace Information Systems*, vol. 13, no. 12, pp. 491–495, 2016. DOI: 10.2514/1.I010432.

[12] C. Bejarano *et al.*, "HARVIS: dynamic rerouting assistant using deep learning techniques for Single Pilot Operations (SPO)," *Transportation Research Procedia*, vol. 66, pp. 262–269, 2022, ISSN: 23521465. DOI: 10.1016/j.trpro.2022.12.026.

[13] Z. T. Zhang, C. Storath, Y. Liu, and A. Butz, "Resilience Through Appropriation: Pilots' View on Complex Decision Support," in *Proceedings of the 28th International Conference on Intelligent User Interfaces*, ser. ACM Digital Library, New York,NY,United States: Association for Computing Machinery, 2023, pp. 397–409, ISBN: 9798400701061. DOI: 10.1145/3581641.3584056.

[14] G. Schmitz and J.-P. Buch, *DLR-Projekt NICo: Vorläufige Auswertung einer Online-Pilotenstudie zur Entscheidungsfindung bei der Auswahl eines Ausweichflughafens*, 2024. DOI: 10.25967/610107.

[15] A. Joshi, S. Kale, S. Chandel, and D. Pal, "Likert Scale: Explored and Explained," *British Journal of Applied Science & Technology*, vol. 7, no. 4, pp. 396–403, 2015. DOI: 10.9734/BJAST/2015/14975.

[16] B. Ceballos, M. T. Lamata, and D. A. Pelta, "A comparative analysis of multi-criteria decision-making methods," *Progress in Artificial Intelligence*, vol. 5, no. 4, pp. 315–322, 2016, ISSN: 2192-6352. DOI: 10.1007/s13748-016-0093-1.

[17] J.-P. Buch and G. Schmitz, *DLR Projekt NICo: Virtuell oder vor Ort? Erkenntnisse über die Onlinedurchführung einer Pilotenstudie zur Entscheidungsfindung*, 2024.

[18] R. Pohl, *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory* (EBL-Schweitzer). London: Taylor and Francis, 2010, ISBN: 9781135844882.

[19] Guillermo Ballester Valor, *Ogimet: Professional information about meteorological conditions in the world*. [Online]. Available: https://www.ogimet.com/ (visited on 08/21/2025).

[20] P. Mayring and T. Fenzl, "Qualitative Inhaltsanalyse," in *Handbuch Methoden der empirischen Sozialforschung*, N. Baur and J. Blasius, Eds., Wiesbaden: Springer Fachmedien Wiesbaden, 2019, pp. 633–648, ISBN: 978-3-658-21307-7. DOI: 10.1007/978-3-658-21308-4_42.

[21] J. E. Birren and L. M. Fisher, "Rules and reason in the forced retirement of commercial airline pilots at age 60," *Ergonomics*, vol. 38, no. 3, pp. 518–525, 1995, ISSN: 0014-0139. DOI: 10.1080/00140139508925123.