

ASSESSING AN ALTERNATE AIRPORT ASSISTANT

J. I. González Cabeza[1], S. Petersen[1], J.-P. Buch[1], D. Niedermeier[1], M. Wies[2] & A. Papenfuß[2]

[1]German Aerospace Centre, Institute of Flight Systems, Lilienthalplatz 7, 38108 Braunschweig, Germany
[2]German Aerospace Centre, Institute of Flight Guidance, Lilienthalplatz 7, 38108 Braunschweig, Germany

Abstract

Recently, industry has worked on concepts for reducing active flight crew in cruise while one crew member is resting (extended Minimum Crew Operations, eMCO). 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) to support and relieve pilots in phases of high workload. As a coping strategy, pilots are trained to adhere to the FOR-DEC decision-making scheme. Mainly to assist with Facts, Options, and Risks categories when considering options for an alternate airport in case of urgent diversions, DLR has designed an electronic flight bag (EFB) application prototype that calculates in-flight and landing performance and provides a ranking of suitable airports. Four equally weighted criteria were used to define suitability through a multi-criteria decision-making approach. These criteria were previously identified as significant in a pilot survey for investigating diversion decisions. The prototype was designed to support the FOR-DEC scheme based on requirements derived from survey findings.

The aims of a simulator study were to assess the functionalities of the EFB application and to gain pilot feedback on its usability and suggestions for further development. Focus was laid on investigating whether the equal weightings are consistent with the pilots' reasoning. Seven crews of two pilots with current or past Airbus A320 type rating performed a scenario in the motion-enabled DLR AVES A320 replica. In this study, most crews decided to divert to an alternate airport when encountering a non-time-critical malfunction in adverse meteorological conditions. The pilots made use of the EFB application during their FOR-DEC process resting importance on its comprehensible ranking mechanism. In an analysis of NASA-TLX and questionnaire results, it became evident that pilots experienced low workload and rated the EFB application favourably. They would consider using it in real operations when faced with the need to divert from the planned destination. Eleven out of 14 pilots indicated they would need less time to decide than without the application. Participants' comments show they approved of it even for today's multi-pilot operations.

Keywords: single pilot operations; electronic flight bag; flight deck; human-machine interface

1. Introduction

With the anticipated substantial increase in global air travel and a shortage of qualified aircrew [1], the concept of reduced crew operations (RCO) has been explored [2]. While economic factors often drive this interest, it is also suggested that human error contributes to many aviation accidents. Proponents of greater automation believe increased use of automation will enhance safety [3]. Recently, industry has been working on strategies to reduce the number of active crew members during cruise when one is resting, known as extended Minimum Crew Operations (eMCO) [4]. Initially planned for 2025 and beyond, eMCO was considered a preliminary step towards single pilot operations (SiPO), albeit EASA effectively halted certification efforts. Hence, neither concept has been certified for EASA CS-25 aircraft.

The NICo project

The German Aerospace Centre's (DLR) Next generation Intelligent Cockpit (NICo) project aimed to create assistance systems for SiPO to aid pilots during high workload periods [5]. Whilst the project did not focus on designing certifiable systems, automation assistance is generally seen as essential for SiPO implementation [6]. These automation systems must reliably support the pilot by managing routine tasks, monitoring aircraft systems, and alerting the pilot to critical situations. According to state-of-the-art reviews, two technological approaches are being explored: one based on existing technologies from unmanned and military aviation, and another relying on higher degrees of autonomy and onboard automation [7, 8]. However, the current technology supporting SiPO concepts remains in prototype form and requires rigorous development to ensure robustness and pilot trust [6]. Currently, pilots in multi-crew settings are trained to follow decision-making frameworks like FOR-DEC. This acronym stands for gathering Facts, generating Options, assessing Risks and benefits, making a Decision, Executing it, and Checking its effectiveness over time [9]. The NICo project introduced the Virtual Co-Pilot (VCP) concept to investigate how automation can help flight crews or single pilots in diagnosing issues and making decisions. Recent expert interviews indicated that pilots would appreciate additional support, especially with the FOR-DEC decision-making process [10]. However, experts stressed that the ultimate decision and its execution should remain with the human crew or pilot, not an automated system. The main objective of developing the VCP was to improve safety in both multi-crew and single-pilot operations, ensuring that all VCP functions are applicable broadly, regardless of the future feasibility of single pilot operations.

Requirements for an alternate airport assistant

Prior to the present study, Schmitz and Buch [11] conducted an online exercise with pilots in two scenarios that made diversions necessary. Participants did not fly in the simulator but performed thought exercises for the two scenarios in a video-conference. They were to imagine that they encountered malfunctions and describe their theoretical courses of action given certain weather and operational conditions. Results included criteria for alternate airports which were rated differently important by participants. From their findings the authors derived requirements for a system to assist upon deciding for an alternate airport. Refer to Table 1 for an excerpt from their ideas for an assistance system concept.

Table 1 – Requirements for an alternate airport assistant derived by Schmitz and Buch [11]

No.	Requirement
1	Allowing for customisation using personally relevant criteria, i.e. weightings according to preferences
2	Not using tables comparing airports based on multiple criteria, as reading tables is time-consuming and prone to errors
3	Excluding airports from assessment that do not fulfil the criteria, especially hard requirements, e.g. stop margin
4	Using failure-specific range calculations
5	Enabling sufficient user interaction while clearly indicating the current settings
6	Allowing pilots to introduce knowledge-based items, e.g. due to experience near the home base
7	Taking into account personal perceptions by numeric filtering, i.e. upper and lower bounds
8	Providing a map with overlays, e.g. highlighting weather
9	Reflecting the FOR-DEC categories in the navigation menu
10	Advising on a rough flight direction

Based upon some of these findings, we have designed an electronic flight bag (EFB) application prototype, mainly to assist with Facts, Options, and Risks categories of FOR-DEC when exploring alternatives for a diversion. The application calculates and evaluates in-flight and landing performance to suggest suitable airports. Four key factors determined suitability of airports: distance,

fuel on arrival, stop margin, and crosswind. Initially, these factors were given equal weight using the multi-criteria decision-making approach TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [12] which is used to compare objects or alternatives based on multiple criteria, where each criterion has potentially a different importance weighting. The method is a widely used and popular approach in many fields, including business, engineering, and environmental science, to support decision-making when multiple criteria are involved. It is relatively easy to implement and can handle both qualitative and quantitative input.

Research hypotheses

Table 2 – Hypotheses regarding workload and EFB application ratings

No.	Hypothesis
1	First officers will experience different workload than captains.
2	Pilots will experience different workload dependent on operational range (i.e. short-range or long-range aircraft).
3	First officers will rate the EFB application differently than captains.
4	Pilots will rate the EFB application differently dependent on operational range (i.e. short-range or long-range aircraft).

Findings from a previous pilot survey confirmed the importance of the four aforementioned criteria in determining where to divert [11]. The aims of the present study were thus to assess the feasibility of the EFB application and to get pilot feedback on its usability. With regard to the pilots using the application, the research hypotheses listed in Table 2 were proposed.
These hypotheses were derived from findings reported previously in literature. It is assumed that captains experience lower workload than first officers during the same scenario because of their additional flight hours. Pilots with more flight hours or recent simulator training exhibited lower physiological stress responses and reported lower perceived workload under the same scenario conditions, indicating that proficiency can mitigate workload spikes [13]. Using the NASA-TLX, another study by Friedrich et al. [14] showed a clear linear relationship between training level and reported workload. Pilots reported lower mental, physical, and temporal demand scores for the same manual-flight scenarios, even when the scenarios were objectively demanding, e.g. due to simulated turbulence, or instrument-failure drills.

2. Methods

Participants

Fourteen pilots, either current or former holders of an Air Transport Pilot Licence (ATPL) and certified to operate the Airbus A320, were enlisted for this research study. The participants had an average age of 49.4 years, with a standard deviation of 12.4 years. On average, they had accumulated 14,740 hours of flight experience, with a standard deviation of 10,870 hours. The group was evenly split, with seven serving as first officers and seven as captains. Categorised by their latest aircraft assignments, five pilots were operating long-range flights, while nine were involved in short-range operations. Where relevant, pilots were compensated for their travel costs but did not receive any additional incentives.

Materials

In the Air Vehicle Simulator (AVES)[1], both pilots were seated in the Airbus A320 replica and given the research application on their electronic flight bags (EFB). This application comprises four primary sections: a status page, a performance page, an airports page, and an alternates page. Each of these pages is readily accessible via the menu bar located at the top of the display.
The alternates page is the most crucial section, offering a ranking of airports that are suitable for diversion (cf. Figure 1). Four criteria determine suitability: distance, fuel upon arrival, stop margin,

[1] AVES is the research simulator of the Institute of Flight Systems located at DLR's Braunschweig site [15].

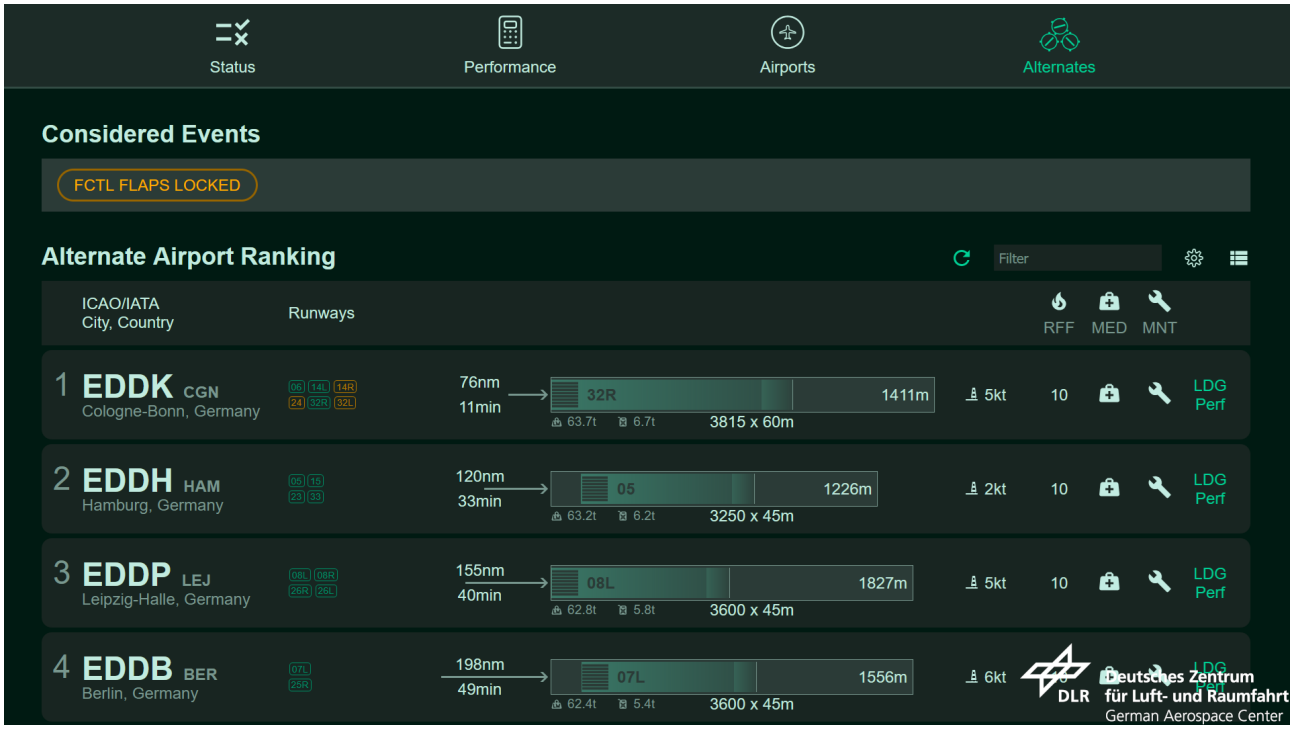


Figure 1 – Alternates page

and crosswind[2], all of which are equally weighted by default using the TOPSIS multi-criteria decision-making method. However, pilots can adjust these weightings based on their experience or preferences. In addition to the airport rankings, essential details such as the dimensions and stop margin of the best available runway, crosswind speed and direction, as well as the remaining distance and time to destination, are displayed. This enables pilots to understand why an airport is ranked higher or lower than others. Hence, this page fulfils the first two requirements outlined by Schmitz and Buch [11], as it allows for user interaction and customisation while showing the current settings through the menu behind the gearwheel icon.

On the page illustrated in Figure 2, performance metrics for a chosen airport are detailed for each runway, allowing pilots to customise settings such as flaps configuration and thrust reverser usage, and to view the calculated landing performance for these settings. In the example seen in the figure, three runways with both directions are depicted. Runways that are insufficiently long are highlighted in red and can be excluded from consideration to disregard options that do not meet the necessary criteria. This page meets the third and fourth requirement.

An airports page depicted in Figure 3 lists all considered airports with their weather conditions. Tapping on an airport entry leads to the performance page for that airport. Pilots thus can retrieve detailed information for an airport that interests them given the forecast weather and runway layout.

Finally, the status page (cf. Figure 4) provides pilots with information about the impact of a detected failure[3] and how it affects the application's calculations, thereby meeting the need for failure-specific details. Additionally, it displays NOTAM[4] and weather information. This page is designed to assist with the Facts category of FOR-DEC by offering crucial information at a glance, thus satisfying the fifth requirement. Moreover, it enhances the understanding of the application's calculations and supports the pilots' mental model.

The core calculations determine the distance and remaining flight time to each airport, along with

[2]For the example of Cologne-Bonn airport in the figure, the airport is 76 nm away, the fuel on arrival is estimated to 6.7 tons, the stop margin for runway 32R is 1411 metres, and the crosswind is 5 knots from the right, i.e. from the north.

[3]The present failure leads to a fuel penalty factor of 1.6 as determined handbook methods and to increased landing distances but does not affect crosswind limits, cf. Figure 4.

[4]Notice to Airmen is filed with an aviation authority and alerts pilots of possible hazards or adverse conditions, e.g. due to construction works.

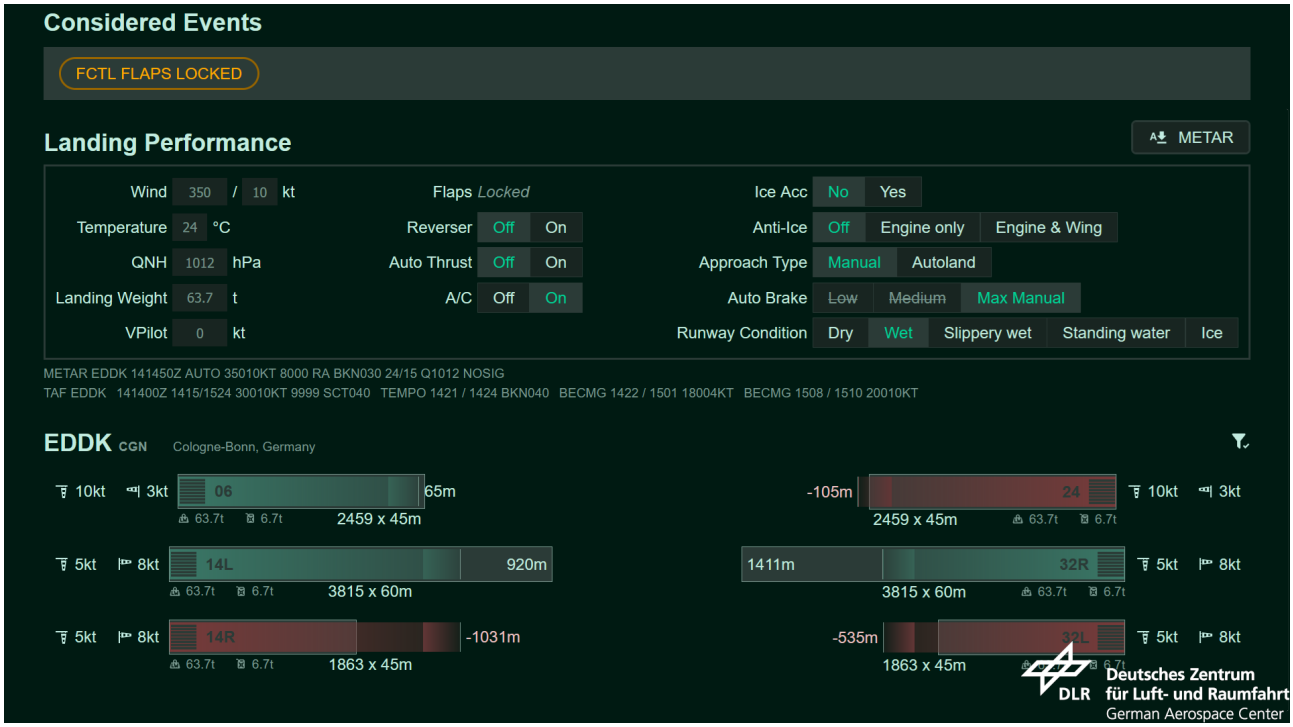


Figure 2 – Performance page

estimates for landing performance, such as the necessary landing distances and speeds, which are calculated taking into account potential failures. The system also considers any anticipated cross-winds and the possibility of increased fuel consumption whose data sources are METAR reports and handbook methods, respectively. Both pilots had access to the application on their individual EFB devices, which were attached to the flight deck panels. Whilst the EFB were not synchronised, they were available for individual use.

Questionnaires

The research utilised several questionnaires, including the NASA-TLX [16] which is a widely used metric to measure subjective workload in various settings. After rating the dimensions, participants also indicate the relative importance of each dimension to the overall task. The final workload score TLX is a weighted average on a sacle from zero to 100. Another questionnaire focused on demographics such as age, gender, nationality, job role, types of aircraft flown, and flight hours. A third covered ergonomics, including situational awareness, trust, and task sharing. Eventually, an application-specific survey (cf. Table 5 in the appendix) addressed aspects of the EFB application.

Procedure

Following a brief familiarisation exercise where each pilot completed a circuit around Frankfurt/Main airport (EDDF) to understand the distinctions between this particular simulator and the actual aircraft, the primary scenario flight took place with both pilots on the simulator flight deck. The pilots decided freely who would take on the role of pilot flying (PF) and who would be the pilot monitoring (PM), i.e. not only captains were PF.

The aircraft was positioned at the holding point of runway 07 at Münster/Osnabrück airport (EDDG), prepared for its flight to Nuremberg (EDDN). Shortly after take-off, as the crew retracted the slats and flaps, they encountered a malfunction where the flaps failed to fully retract and got stuck between positions 0 and 1. This issue resulted in increased fuel consumption due to greater drag and necessitated a longer landing distance. With adverse weather affecting several aerodromes, the pilots had to carefully decide on an alternate airport. Most pilots diverted from the planned destination EDDN, whilst the study design was such that a return to EDDG was impossible for the remaining runway length being insufficient in present meteorological conditions.

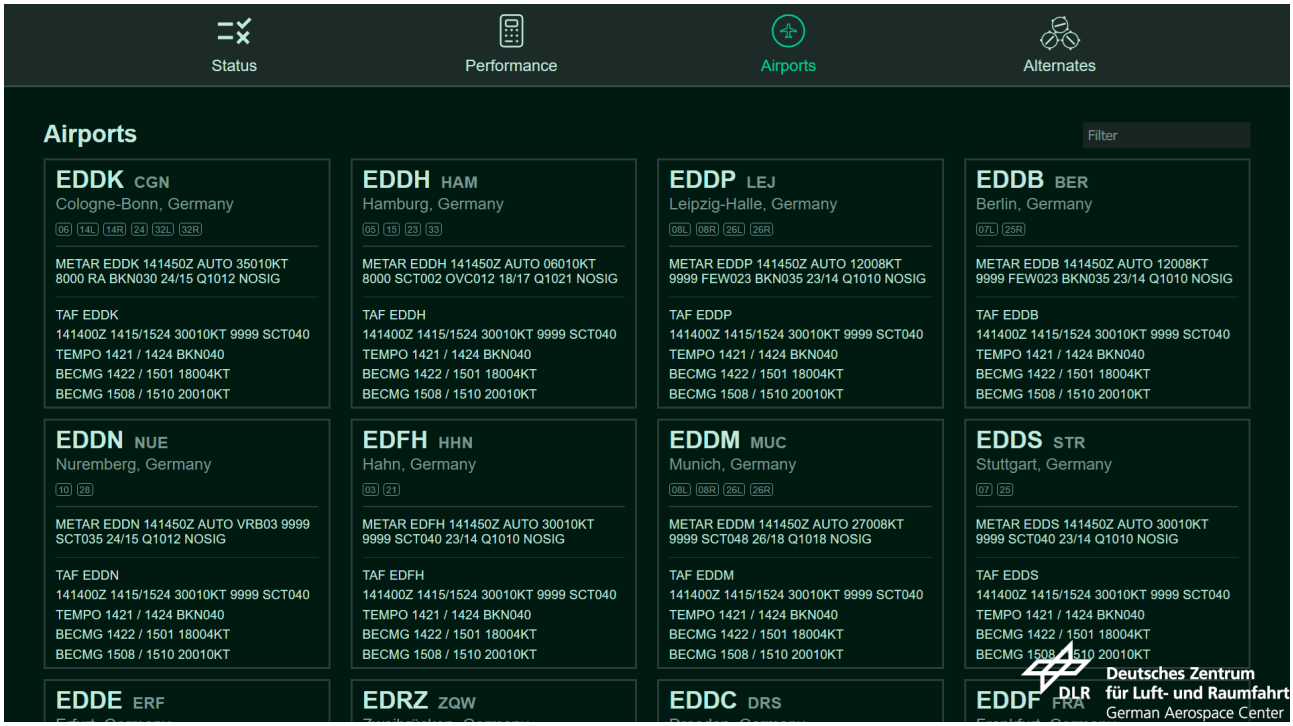


Figure 3 – Airports page

Throughout the scenario flight, recordings were made of simulator data, as well as video and audio. The simulator data encompassed flight dynamics parameters and aircraft system states, while the video and audio captured the flight deck view, instrument broadcasts, and interactions among the flight crew, simulated ATC, and researchers.

3. Results

This section presents results of the NASA-TLX instrument and the EFB-specific questionnaire. To answer the research hyptoheses, Mann-Whitney U tests were employed. These assess whether two independent groups come from the same distribution. They are a nonparametric alternative to the independent-samples t-test, useful when the data are ordinal, or not normally distributed. U denotes the test statistic and n the sample size of each group. The null hypothesis, i.e. that the two distributions are identical in location, is commonly rejected if the p-value is less than 0.05. Pilots were grouped by operations (ops), i.e. short-range (S) or long-range (L) aircraft, and by position (pos), i.e. first officer (F) or captain (C).

NASA-TLX

The NASA-TLX questionnaire was administered upon concluding the experiment. Figure 5 shows the weighted scores of twelve of the subjects across the dimensions mental demand (MD), physical demand (PD), temporal demand (TD), performance (PE), effort (EF), and frustration (FR), as well as the aggregate task load index (TLX). On average, the pilots indicated a TLX of 7.41 which is fairly low an a scale from zero to 100. Partial scores of mental demand were the highest and those of frustration the lowest. To test for differences between pilot position and operational type regarding the TLX scores and thus to answer the first two hypotheses, a series of Mann-Whitney U tests was employed. There was no significant difference, i.e. p was above 0.05 for all tests, concerning the ranks of the items on the TLX scales, refer to Table 3 for the results.

EFB-specific questionnaire

Descriptive statistics

A questionnaire comprising 47 items (cf. Table 5 in the appendix) was administered to measure pilots' agreement with statements regarding the different pages and features of the EFB application.

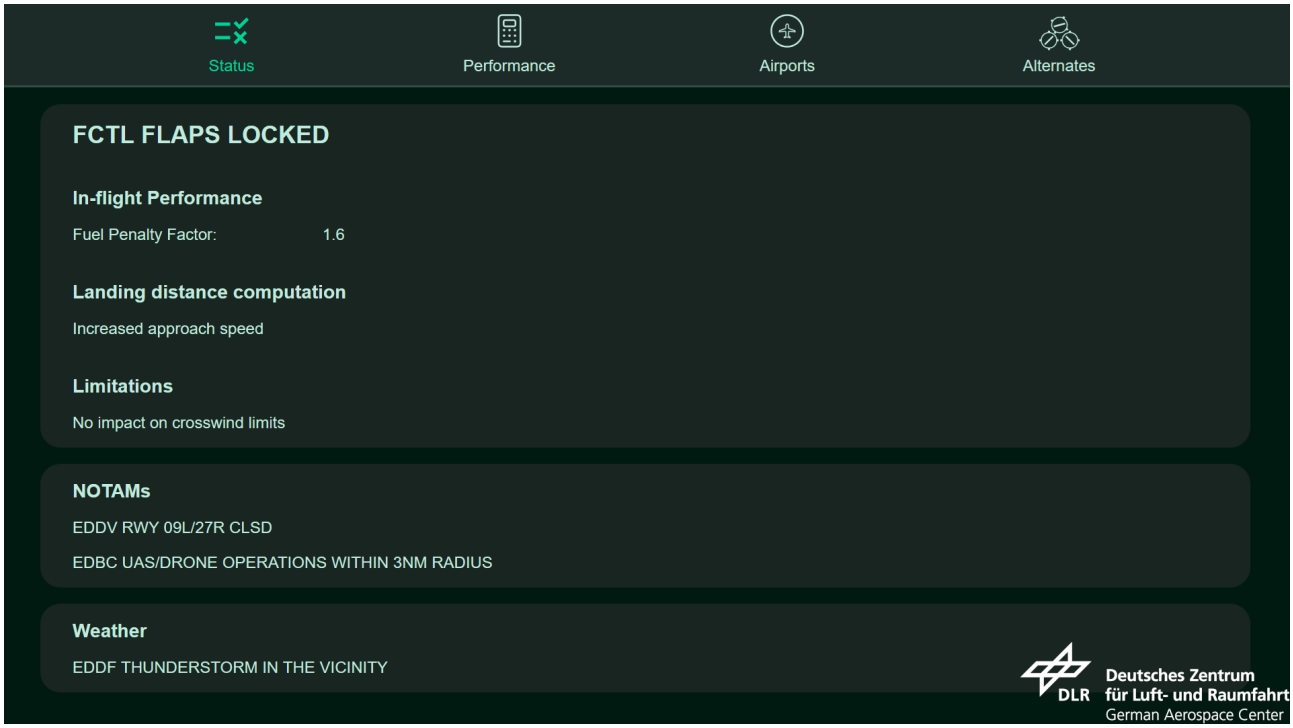


Figure 4 – Status page

We used four-point Likert scales, ranging from “totally disagree” (1) to “totally agree” (4). Hence, no neutral response was possible. Table 6 in the appendix lists median, interquartile range (IQR), minimum and maximum, as well as percentiles for each of the items. Agreement was observed for most questionnaire items, i.e. the median response often was at least 3 corresponding to “I somewhat agree with the statement”. Interquartile ranges as measurement of dispersion were small for most items suggesting low variability in the data.

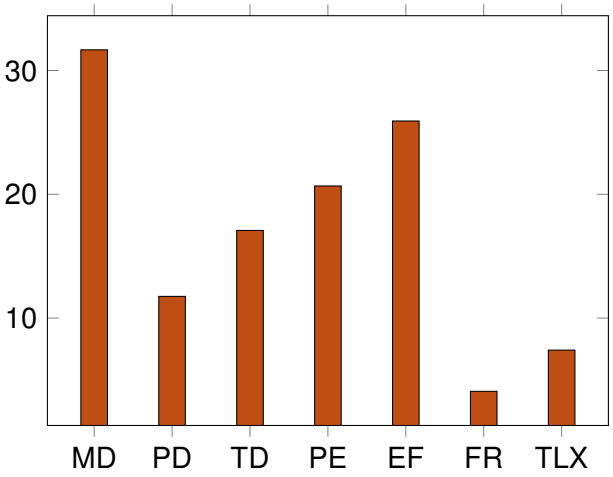


Figure 5 – NASA-TLX average weighted scores, reprinted from [5]

Table 3: Independent Samples Mann-Whitney U Tests across operations and position for TLX scores

TLX dimension	U_{ops}	n_S	n_L	p_{ops}	U_{pos}	n_F	n_C	p_{pos}
Mental Demand	15.50	9	5	.807	12.50	7	7	.422
Physical Demand	20.00	9	5	.743	17.50	7	7	1.000
Temporal Demand	14.00	9	5	.622	17.00	7	7	.936
Performance	12.00	9	5	.415	15.50	7	7	.748
Effort	11.00	9	5	.329	21.00	7	7	.688
Frustration	13.00	9	5	.440	17.00	7	7	.924
TLX	13.50	9	5	.569	16.00	7	7	.810

L = long-range, S = short range

C = captain, F = first officer

All pages, even the status page which contained only little information, were rated favourably, receiving a median of at least 3 ("I somewhat agree with the statement"). Key aspects of this study were the performance and alternate pages. These were rated most helpful by the pilots, especially with the alternates page receiving a median of 4 ("I totally agree"). Figure 6 depicts the corresponding response frequencies of the different ratings. Not all participants answered every item.

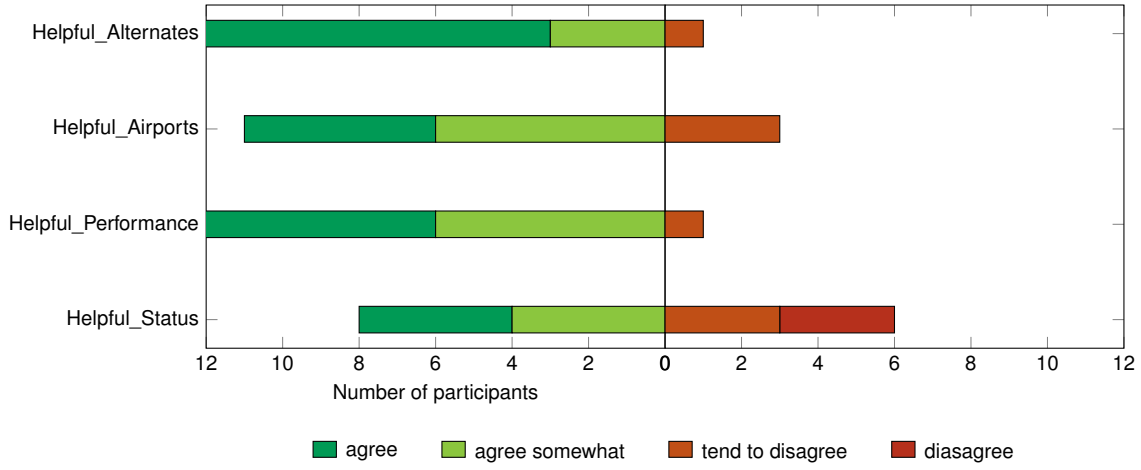


Figure 6 – Response distribution comparing pages' helpfulness; not all participants answered every item

Hypothesis tests

In order to test the third and fourth hypothesis concerning the responses to the EFB-specific questionnaire, another series of Mann-Whitney U tests was employed to test for differences in-between pilots. In many comparisons there was no significant difference concerning the ranks of the items on the Likert scales, except for five items.

Refer to Table 7 in the appendix for the results, while Table 4 lists only the significant items. H_{ops} and H_{pos} denote which of the distributions ranked higher than the other. Noteworthy results included the statement that the customisable weightings were useful which was significantly more endorsed by short-range pilots. Another example is the map display to be implemented that was notably higher ranked by first officers.

Table 4: Independent Samples Mann-Whitney U Tests across operations and position for EFB-specific questionnaire items with significant differences.
Bold font denotes items with $p < .05$

Item ID	U_{ops}	U_{ops}	n_S	n_L	p_{ops}	H_{ops}	U_{pos}	n_F	n_C	p_{pos}	H_{pos}
Status_Inflightperf	7.5	9	5	**0.036**	S	21	7	7	0.677		
Status_Limitations	7.5	9	5	**0.036**	S	32.5	7	7	0.299		
Alternates_CustomWeights	4.5	9	5	**0.011**	S	22	7	7	0.781		
Alternates_FlightTime	16.5	9	5	0.425		9	7	7	**0.037**	F	
Statement_Map	21.5	9	5	0.941		9	7	7	**0.032**	F	

H_{ops}: distribution with higher rank; L = long-range, S = short range
H_{pos}: distribution with higher rank; C = captain, F = first officer

Selected statements in detail

Refer to Figure 7 for the response frequencies for another subset of the questionnaire items. Again, not all participants answered every item. Eleven pilots indicated that in a diversion scenario they would need less time to decide than without the application. Eleven pilots at least somewhat agreed to the statement that the underlying calculations are comprehensible. Moreover, eleven subjects indicated that they agreed to the statement all criteria necessary for decision-making are included. Lastly, nine pilots at least somewhat agreed to the statement the EFB application strongly influenced their decision-making. Still, 13 pilots affirmed their actual decision would eventually be the same in both cases.

Pilot feedback

Upon concluding the experiment, many participants gave valuable feedback on functions and user interface of the application, including suggestions to consider the current runway in use, weather-dependent colour-coding, customisable minima and maxima for the airport criteria, and to display the approach speed.
Whilst pilots admitted to need more time to get used to the application, they affirmed it would perfectly support FOR-DEC. It could be further improved by highlighting restricted airport options, changing the wind information icon, and considering the actual runway condition reports.

4. Discussion

Pilot workload

The research aimed to assess the feasibility of a newly developed EFB application that should support pilots decision-making when considering options where to divert. Pilots experienced low workload (average TLX score 7.41) throughout the scenario and were on average neither highly demanded nor frustrated. This supports the statement that the EFB application is a tool helpful to them, speeding up their established processes.
For a different scenario where pilots operated spatially separated and without the EFB application, Niedermeier et al. [17] reported that pilots felt rather overwhelmed and not being able to deliver their full performance. The authors believe that our application would be especially useful in such demanding scenarios, or in an eMCO setting, where it could help the awakening pilot.
Contrary to findings in literature [13, 14], in our study workload ratings did not vary in function of operations and position. The authors presume this is due to the fact that overall workload ratings were low as we deliberately did not put participants in a time critical situation.

Application ratings

The questionnaire results revealed that pilots rated the application agreeably and found its features useful, particularly the alternates ranking which is based on a multi-criteria decision-making method. The corresponding page was rated most helpful amongst all. The research hypotheses however, could just be partially supported by the results. Some differences in application ratings were evident regarding the operational range (i.e. short-range or long-range flights), and the pilot position (i.e.

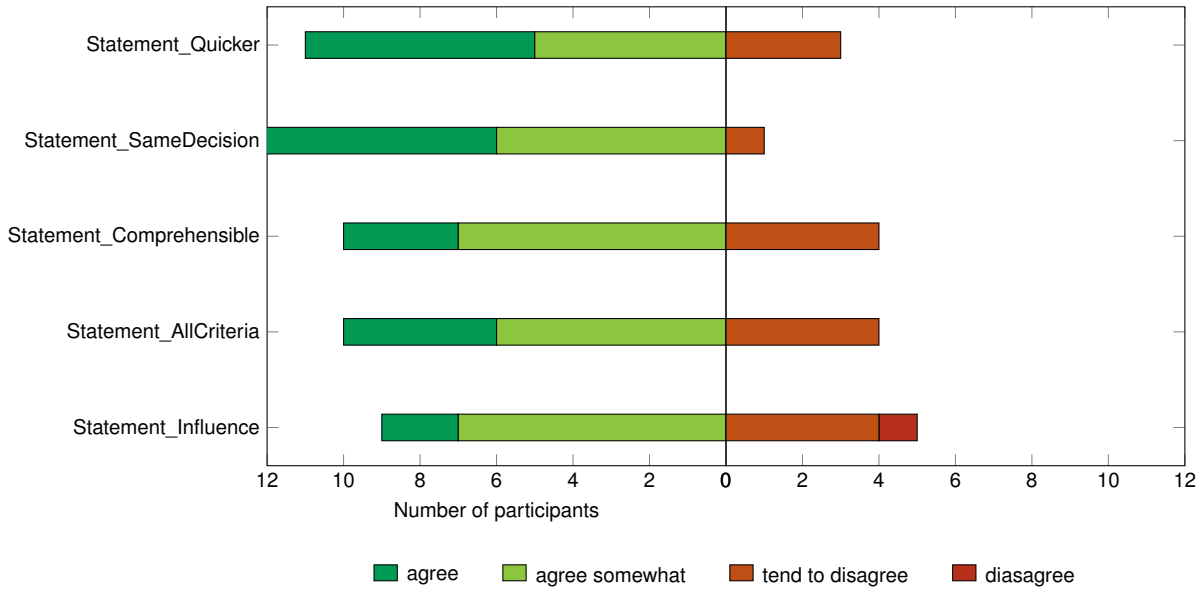


Figure 7 – Response distribution for selected questionnaire items; not all participants answered every item

captain or first officer) which were significant for five of the 47 items. We conclude that the application would be of use for both short- and long-range flights and regardless of the pilot position, albeit some features might benefit especially first officers or those operating short-haul flights.
Eleven pilots indicated that in a diversion scenario they would decide quicker than without the EFB-application. This shows that our application works as intended. It supports the pilots in their decision-making process by providing information in a helpful way. Considering the pilots coming eventually to the same decision, it speeds up the process without distorting the results which is important, especially in time-critical situations.
Pilots rested importance on the comprehensibility of the ranking mechanism. They affirmed to have understood why an airport was ranked higher or lower, respectively. The customisable weightings of the ranking and filtering of airports allowed for pilots' personal preferences to be considered. The latter of the two features, however, was not deemed useful by participants. This is not in line with the requirement no. 7 postulated by Schmitz and Buch [11] who explicitly suggested numerical filtering to be beneficial. We presume this may be true for some pilots but also be a matter of personal taste. Whilst some pilots explicitly stated they would prefer a map displayed, answers to the respective questionnaire item rather indicated the opposite. With these mixed responses, a map might be a nice-to-have or requirement no. 8 might even be obsolete.

Reliability

While designing a certifiable system was not the aim of this study, the reliability of the application needs to be ensured in future development efforts. Pilots are assumed to rely on the EFB application, as nine pilots indicated being at least somewhat influenced by it in their decision-making. This is an interesting finding that raises the question of overtrust. Reliance on assistance systems can lead to challenges such as overtrust or complacency, potentially diminishing manual flying skills and critical thinking in unexpected situations. Effective decision-making thus depends on pilots ability to appropriately interpret and integrate assistance system outputs while maintaining active engagement and situational judgment. Pilot overtrust in automated systems occurs when pilots rely too heavily on automation, reducing vigilance and situational awareness [18].
The authors believe that overtrust was not an issue here as pilots' comments indicated they used the application confidently and had more time to look at primary instruments thus fostering vigilance. Future efforts should maintain appropriate trust levels and ensure a more robust system eventually being certifiable.

Limitations and future work

Researchers are encouraged to consider pilot feedback obtained when further developing the EFB application and to continue involving pilots in the design process. They commented on functionalities and user interface indicating they were very content with the application. Apart from constructive criticism, they approved of it supporting FOR-DEC. Present study used a prototype that was designed to test functionalities rather than user interface design. Consequently, forthcoming research is needed to identify learnability and efficiency, i.e. the time new users need to comprehend the application, and the performance of more experienced users, respectively.
All results need to be interpreted within the confines of the research. Participants were self-selected and the sample of 14 pilots was relatively small. This limited number of subjects does not allow any generalisation of the study results.

5. Conclusion

Within the project NICo, an exploratory simulator study was performed to assess the feasibility of a newly-developed EFB application in a failure scenario. Special emphasis was laid on the comprehensibleness of criteria considered in the alternate airport ranking. Although only seven crews took part in the experiments, interesting qualitative results based on questionnaires and crew comments were obtained. The questionnaire responses indicate that the features of the application are valuable upon decision-making. Crew comments showed the application's potential for use in commercial operations. The EFB application could support pilots in common multi-crew configurations or help the awakening pilot in an eMCO scenario shape their mental picture.

6. Corresponding Author

To reach the corresponding author, mail to inaki.gonzalezcabeza@dlr.de or call +49 531 295 1181.

7. Copyright Statement

The authors confirm that they, and/or their company or organisation, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

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.

[3] A. A. Ylmaz, "Flight in Transition: Navigating the Skies of Automation and Human Judgement," *Journal of Aviation*, vol. 8, no. 3, pp. 305–314, 2024. DOI: 10.30518/jav.1486331.

[4] 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.

[5] J. I. GONZÁLEZ CABEZA, S. Petersen, J.-P. Buch, D. Niedermeier, M. Wies, and A. Papenfuss, "DLR Project NICo – Feasibility Study of an Alternate Airport Assistant for Pilots Deciding Where to Divert," in *CEAS - AIDAA Conference 2025*, ser. Materials Research Proceedings, Materials Research Forum LLC, 2026, pp. 1471–1474. DOI: 10.21741/9781644904251-258.

[6] K.-P. L. Vu, J. Lachter, V. Battiste, and T. Z. Strybel, "Single Pilot Operations in Domestic Commercial Aviation," *Human factors*, vol. 60, no. 6, pp. 755–762, 2018. DOI: 10.1177/0018720818791372.

[7] D. Harris, "Single-pilot airline operations: Designing the aircraft may be the easy part," *The Aeronautical Journal*, vol. 127, no. 1313, pp. 1171–1191, 2023, ISSN: 0001-9240. DOI: 10.1017/aer.2022.110.

[8] N. Puca and G. Guglieri, "Enabling Civil Single-Pilot Operations: A State-of-the-Art Review," *Aerotecnica Missili & Spazio*, vol. 104, no. 3, pp. 187–212, 2025, ISSN: 0365-7442. DOI: 10.1007/s42496-024-00223-7.

[9] H. J. Hörmann, Ed., *FOR-DEC - A Prescriptive Model for Aeronautical Decision Making*, 1994. [Online]. Available: https://elib.dlr.de/27044/.

[10] 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/.

[11] G. Schmitz and J.-P. Buch, "DLR project NICo: findings of an online pilot study on decision-making with regard to the selection of an alternate airport," *CEAS Aeronautical Journal*, 2026, ISSN: 1869-5582. DOI: 10.1007/s13272-025-00935-x.

[12] B. Uzun, M. Taiwo, A. Syidanova, and D. Uzun Ozsahin, "The Technique For Order of Preference by Similarity to Ideal Solution (TOPSIS)," in *Application of Multi-Criteria Decision Analysis in Environmental and Civil Engineering*, ser. Professional Practice in Earth Sciences, D. Uzun Ozsahin, H. Gökçeku, B. Uzun, and J. LaMoreaux, Eds., Cham: Springer International Publishing, 2021, pp. 25–30, ISBN: 978-3-030-64764-3. DOI: 10.1007/978-3-030-64765-0{\textunderscore}4.

[13] A. Faulhaber, "From Crewed to Single-Pilot Operations: Pilot Performance and Workload Management," *20th International Symposium on Aviation Psychology*, pp. 283–288, 2019. [Online]. Available: https://corescholar.libraries.wright.edu/isap_2019/48.

[14] M. Friedrich, S. Y. Lee, P. Bates, W. Martin, and A. K. Faulhaber, "The influence of training level on manual flight in connection to performance, scan pattern, and task load," *Cognition, Technology & Work*, vol. 23, no. 4, pp. 715–730, 2021, ISSN: 1435-5558. DOI: 10.1007/s10111-020-00663-8.

[15] H. Duda, S. K. Advani, and M. Potter, "Design of the DLR AVES Research Flight Simulator," in *AIAA Modeling and Simulation Conference Proceedings*, AIAA, Ed., 2013. DOI: 10.2514/6.2013-4737.

[16] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," in *Human Mental Workload*, ser. Advances in Psychology, vol. 52, Elsevier, 1988, pp. 139–183, ISBN: 9780444703880. DOI: 10.1016/S0166-4115{\%}2808{\%}2962386-9.

[17] 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.

[18] M. Causse, M. Mercier, O. Lefrançois, and N. Matton, "Impact of automation level on airline pilots' flying performance and visual scanning strategies: A full flight simulator study," *Applied ergonomics*, vol. 125, p. 104 456, 2025. DOI: 10.1016/j.apergo.2024.104456.

A EFB-specific questionnaire

Table 5: EFB-specific questionnaire items (translated from the German original)

Group	Statement	Item ID
Please indicate how much each page of the app helped you in your decision-making	The Status page helped me	Helpful_Status
	The Performance page helped me	Helpful_Performance
	The Airports page helped me	Helpful_Airports
	The Alternates page helped me	Helpful_Alternates
Please indicate how intuitive each of the apps pages was to use	The Status page was intuitive to use	Intuitive_Status
	The Performance page was intuitive to use	Intuitive_Performance
	The Airports page was intuitive to use	Intuitive_Airports
	The Alternates page was intuitive to use	Intuitive_Alternates
Please indicate how useful you found the respective information on the Status page	The error message was useful	Status_Error
	The In-flight Performance note was useful	Status_Inflightperf
	The Landing Distance Computation note was useful	Status_Landingdistance
	The Limitations notice was useful	Status_Limitations
	The NOTAM notice was useful	Status_NOTAM
	The Weather notice was useful	Status_Weather
Please indicate how useful you found the respective information on the Performance page	Landing Performance input fields were useful	Peformance_Landing-perf
	METAR/TAF texts were useful	Performance_METAR
	Crosswind windsock was useful	Performance_Xwindsock
	Headwind windsock was useful	Performance_Hwindsock
	Runway display was useful	Performance_RWY
	Filter for runways that are too short was useful	Performance_shortRWYs
Please indicate how useful you found the respective information on the Airports page	The filter/search function was useful	Airports_Filter
	Runway identifiers were useful	Airports_RWYidentifier
	METAR/TAF texts were useful	Airports_METAR
	Sorting of airports was useful	Airports_Sorting

Continued on next page

Table 5: EFB-specific questionnaire items (translated from the German original) (Continued)

Group	Statement	Item ID
Please indicate how useful you found the respective information on the Alternates page	Filter/search function was useful	Alternates_Filter
	Customisable weightings were useful	Alternates_CustomWeights
	Toggling additional information was useful	Alternates_SwitchInfo
	Colour-coding of runway identifiers was useful	Alternates_Colourcoding
	Distance information was useful	Alternates_Distance
	Flight time information was useful	Alternates_FlightTime
	The runway diagram was useful	Alternates_RWY
	The crosswind wind sock was useful	Alternates_Xwindsock
	The RFF category was useful	Alternates_RFF
	The MED category was useful	Alternates_MED
	The MNT category was useful	Alternates_MNT
Please indicate to what extent you agree with the following statements regarding the Alternates page	It was clear to me why an airport is listed higher or lower on the Alternates page	Statement_Ranking
	I would like runway width to be included as an additional weighting criterion	Statement_Width
	I would like the tailwind component to be an additional weighting criterion	Statement_Tailwind
	I would like the cloud base to be an additional weighting criterion	Statement_Ceiling
	I would also like all NOTAMs to be displayed	Statement_NOTAM
	I would like to see numerical filters, e	Statement_Numerical
	I would also like to see a map showing the airports	Statement_Map
Please indicate to what extent you agree with the following statements regarding the app in general	The app has helped me reach a decision on an airport more quickly	Statement_Quicker
	Without the app, I would have reached the same decision regarding an airport	Statement_SameDecision
	The values underlying the calculation are transparent	Statement_Reasonable
	All criteria necessary for a decision are included	Statement_AllCriteria
	The app has strongly influenced my decision-making	Statement_Influence

B Descriptive statistics

Table 6: Descriptive statistics for EFB-specific questionnaire items

Item ID	Median	IQR	Minimum	Maximum	25th p.	50th p.	75th p.
Helpful_Status	3.000	1.750	1.000	4.000	2.000	3.000	3.750
Helpful_Performance	3.500	1.000	2.000	4.000	3.000	3.500	4.000
Helpful_Airports	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Helpful_Alternates	4.000	0.750	2.000	4.000	3.250	4.000	4.000
Intuitive_Status	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Intuitive_Performance	3.000	1.000	3.000	4.000	3.000	3.000	4.000
Intuitive_Airports	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Intuitive_Alternates	3.000	1.000	3.000	4.000	3.000	3.000	4.000
Status_Error	4.000	1.750	1.000	4.000	2.250	4.000	4.000
Status_Inflightperf	2.500	2.000	1.000	4.000	2.000	2.500	4.000
Status_Landingdistance	3.500	1.000	2.000	4.000	3.000	3.500	4.000
Status_Limitations	3.500	2.000	1.000	4.000	2.000	3.500	4.000
Status_NOTAM	3.000	1.000	1.000	4.000	3.000	3.000	4.000
Status_Weather	3.000	1.750	1.000	4.000	2.250	3.000	4.000
Peformance_Landingperf	4.000	1.000	2.000	4.000	3.000	4.000	4.000
Performance_METAR	4.000	1.000	2.000	4.000	3.000	4.000	4.000
Performance_Xwindsock	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Performance_Hwindsock	3.000	1.750	2.000	4.000	2.250	3.000	4.000
Performance_RWY	4.000	0.000	3.000	4.000	4.000	4.000	4.000
Performance_shortRWYs	4.000	0.750	2.000	4.000	3.250	4.000	4.000
Airports_Filter	2.000	1.000	1.000	4.000	2.000	2.000	3.000
Airports_RWYidentifier	4.000	0.750	2.000	4.000	3.250	4.000	4.000
Airports_METAR	4.000	1.000	3.000	4.000	3.000	4.000	4.000
Airports_Sorting	3.000	1.750	1.000	4.000	2.250	3.000	4.000
Alternates_Filter	2.500	1.000	1.000	4.000	2.000	2.500	3.000
Alternates_CustomWeights	3.500	1.750	2.000	4.000	2.250	3.500	4.000
Alternates_SwitchInfo	2.000	1.000	2.000	4.000	2.000	2.000	3.000
Alternates_Colourcoding	4.000	1.000	3.000	4.000	3.000	4.000	4.000
Alternates_Distance	4.000	0.000	3.000	4.000	4.000	4.000	4.000
Alternates_FlightTime	3.500	1.000	2.000	4.000	3.000	3.500	4.000
Alternates_RWY	4.000	0.000	3.000	4.000	4.000	4.000	4.000
Alternates_Xwindsock	4.000	1.000	2.000	4.000	3.000	4.000	4.000
Alternates_RFF	3.000	0.750	2.000	4.000	2.250	3.000	3.000
Alternates_MED	3.000	0.750	2.000	4.000	3.000	3.000	3.750
Alternates_MNT	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Statement_Ranking	4.000	1.000	2.000	4.000	3.000	4.000	4.000
Statement_Width	3.000	1.000	2.000	4.000	2.000	3.000	3.000
Statement_Tailwind	3.000	1.000	2.000	3.000	2.000	3.000	3.000
Statement_Ceiling	2.500	2.000	2.000	4.000	2.000	2.500	4.000
Statement_NOTAM	2.000	0.750	1.000	4.000	2.000	2.000	2.750
Statement_Numerical	3.000	1.750	2.000	4.000	2.000	3.000	3.750
Statement_Map	2.000	1.000	2.000	4.000	2.000	2.000	3.000

Continued on next page

Table 6: Descriptive statistics for EFB-specific questionnaire items (Continued)

Item ID	Median	IQR	Minimum	Maximum	25th p.	50th p.	75th p.
Statement_Quicker	3.000	1.000	2.000	4.000	3.000	3.000	4.000
Statement_SameDecision	3.500	1.000	2.000	4.000	3.000	3.500	4.000
Statement_Reasonable	3.000	0.750	2.000	4.000	2.250	3.000	3.000
Statement_AllCriteria	3.000	1.500	2.000	4.000	2.250	3.000	3.750
Statement_Influence	3.000	1.000	1.000	4.000	2.000	3.000	3.000

C Mann-Whitney U Tests for EFB-specific questionnaire

Table 7: Independent Samples Mann-Whitney U Tests across operations and position for EFB-specific questionnaire items.
Bold font denotes items with p < .05

Item ID	U_{ops}	df_{ops}	p_{ops}	U_{pos}	df_{pos}	p_{pos}
Helpful_Status	15		0.335	33.5		0.262
Helpful_Performance	17		0.456	29.5		0.52
Helpful_Airports	18.5		0.617	20.5		0.632
Helpful_Alternates	19.5		0.674	17		0.259
Intuitive_Status	14		0.253	21		0.681
Intuitive_Performance	21.5		0.938	31.5		0.334
Intuitive_Airports	19		0.655	22.5		0.83
Intuitive_Alternates	21.5		0.938	31.5		0.334
Status_Error	20.5		0.824	25.5		0.943
Status_Inflightperf	7.5		**0.036**	21		0.677
Status_Landingdistance	11		0.111	29		0.578
Status_Limitations	7.5		**0.036**	32.5		0.299
Status_NOTAM	12		0.154	32.5		0.306
Status_Weather	10		0.09	34		0.223
Peformance_Landingperf	13		0.156	35.5		0.112
Performance_METAR	16.5		0.41	30.5		0.429
Performance_Xwindsock	12.5		0.175	21		0.681
Performance_Hwindsock	14.5		0.286	33.5		0.247
Performance_RWY	23		1	21		0.591
Performance_shortRWYs	26		0.614	17		0.259
Airports_Filter	13		0.207	21.5		0.737
Airports_RWYidentifier	13		0.13	32		0.259
Airports_METAR	28		0.423	35		0.125
Airports_Sorting	20		0.779	18		0.42
Alternates_Filter	11		0.127	19.5		0.55
Alternates_CustomWeights	4.5		**0.011**	22		0.781
Alternates_SwitchInfo	15.5		0.309	34		0.177
Alternates_Colourcoding	16.5		0.393	31.5		0.334
Alternates_Distance	NaN			NaN		
Alternates_FlightTime	16.5		0.425	9		**0.037**
Alternates_RWY	20.5		0.742	24.5		1

Continued on next page

Table 7: Independent Samples Mann-Whitney U Tests across operations and position for EFB-specific questionnaire items.
Bold font denotes items with p < .05 (Continued)

Item ID	U_{ops}	df_{ops}	p_{ops}	U_{pos}	df_{pos}	p_{pos}
Alternates_Xwindsock	14.5		0.24	21.5		0.708
Alternates_RFF	13.5		0.218	11.5		0.083
Alternates_MED	21		0.885	16.5		0.298
Alternates_MNT	14		0.253	14		0.171
Statement_Ranking	19.5		0.708	27.5		0.72
Statement_Width	24		0.879	25.5		0.942
Statement_Tailwind	21		0.873	35		0.125
Statement_Ceiling	13		0.188	20.5		0.624
Statement_NOTAM	16		0.348	20		0.549
Statement_Numerical	24.5		0.832	26.5		0.839
Statement_Map	21.5		0.941	9		**0.032**
Statement_Quicker	22		1	17		0.338
Statement_SameDecision	33.5		0.118	26		0.886
Statement_Reasonable	19		0.664	32.5		0.298
Statement_AllCriteria	17.5		0.522	34.5		0.195
Statement_Influence	18		0.564	16		0.268