Human Error, Reliability, Resilience, and Performance, Vol. 82, 2023, 21–32
https://doi.org/10.54941/ahfe1003550

Human Reliability Analysis in Aviation Accidents: A Review

Steven Tze Fung Lam[1,2] and Alan Hoi Shou Chan[2]

[1]Industrial Centre, The Hong Kong Polytechnic University, Hong Kong SAR, China
[2]Department of Advanced Design and Systems Engineering, City University of Hong Kong, Hong Kong SAR, China

ABSTRACT

In the civil aviation sector, human factors is the primary cause of many safety incidents. Aircraft flying, maintenance, and operations are the major tasks that are heavily dependent on professionals; thus, they are subject to human error probability. Human reliability analysis (HRA), which can evaluate human state and managing risk, has been developed over the years to identify, predict, and reduce human failures throughout aircraft operating procedures. Different generations of HRA tools have been developed to quantify the risks that are associated with safety accidents, including such as the Human Error Assessment and Reduction Technique, Technique for Human Error Prediction, Standardized Plant Analysis Risk Human Reliability Analysis, Cognitive Reliability and Error Analysis Methods, and Bayesian Network (BN). However, little is known about how these approaches are applied in aviation safety. This review aimed to systematically examine the current status of research on HRA in aviation accidents. A total of 13 studies were included and encompassed the studies of the first, second, and third generalizations of HRA alone or in combination with other methods (e.g., Improved Analytics Hierarchy Process, Functional Resonance Analysis Methods, Human Factor Analysis and Classification System, and Fault Tree Analysis). Results revealed that the third-generation HRA with BN was frequently used, showing great application potential for flight safety risk prevention and reduction. In the future, testing other third-generation HRA models driven by data in the field of airworthiness is necessary.

Keywords: Aviation, Safety, Human reliability analysis, Performance shaping factors, Human error

INTRODUCTION

Despite the advancement in aircraft technology and increased awareness of aviation safety, airplane accidents remain frequently reported in aircraft airliners (Transportation Safety Board of Canada, 2023). Aircraft accidents are defined as "the events occurring due to non-functionality of the operations in the aircraft from the time of boarding of the passengers till the landing of the flights where all the passengers have disembarked" (Abeyratne & Abeyratne, 2012). According to the international Civil Aviation Organization, 60-80% of aviation accidents are attributed to human errors, particularly pertaining to pilots (Boyd, 2017; Erjavac, Lammartino, & Fossaceca, 2018). By definition, human error is 'the failure of planned actions to achieve their desired ends-without the intervention of some unforeseeable event' (Reason, 2016). In preventing and reducing human error, human reliability analysis (HRA) involves a structured process to determine and estimate the probability of human failure events. In general, HRA covers the following: recognizing human errors associated with specific tasks, modelling critical human errors, and calculating the human error probability (HEP) (Pan, Lin & He, 2017).

According to a number of reviews, different generations of HRA tools have been developed to determine the risks that are associated with safety accidents, including the first-generation HRA (e.g., Human Error Assessment and Reduction Technique (HEART) and the Technique for Human Error Prediction [THERP]), second-generation HRA (e.g., Standardized Plant Analysis Risk Human Reliability Analysis [SPAR-H] and Cognitive Reliability and Error Analysis Methods [CREAM]), and third-generation HRA (e.g., Bayesian Network [BN] and Phoenix) (Hou, et al., 2021; Tao, Yang, and Duan, 2020). However, little is known about how these generations of HRA are applied in aviation safety. Despite HRA has been widely used in safety critical domains such as nuclear power, oil and gas, it is relatively less adopted in the aviation field because the Human Factors Analysis and Classification System (HFACS) method predominated in the analysis of aviation accidents (Wiegmann and Shappell, 2003; Mkrtchyan, et al., 2015; Wan, et al., 2020). Hence, the aim of this review was to systematically examine the current status of research on HRA in aviation accidents. The results of this review would provide practitioners and researchers with valuable information on theoretical and practical applications of HRA in aviation and identify potential areas for improvement, including training, procedures, and equipment design.

METHODS

This literature review was conducted through the Web of Science database. This database was selected because it is mostly used for retrieving highly impactful research articles. In maximizing the search results, two sets of keywords were used in combination. The first set of keywords included "human reliability analysis," human error probability," "human failure event," "human error analysis," "probabilistic risk assessment," "operator error," "human performance factor," "performance influencing factor," "performance shaping factor," and "situation awareness." The second set of keywords included "aviation safety," "aviation risk," "flight safety," "flight operational quality assurance," "general aviation," "aviation accidents," and "flight accidents."

No time limit was set for the search because this is the first review on HRA in the aviation safety sector. The search date was 26 December 2022. Articles should meet the inclusion criteria: (a) used either the first, second, and third-generation HRA for identification and analysis of human errors; (b) focused on aviation safety-related flight crew tasks, at least one flight operation task; and (c) published in English language peer-reviewed articles or conference proceedings. The exclusion criteria were as follows: (a) introduction of any generation of HRA; (b) analysis using the Human Factors Analysis and Classification System (HFACS) alone; and (c) case studies, textbooks, and dissertations.

After data entry into EndNote 20, duplicates were eliminated. Authors checked each record based on title and abstract following inclusion and exclusion criteria, and removed ineligible records. This was followed by full-text screening. Excluded records were provided with reasons. Two authors independently screened the records to identify relevant articles. Any disagreements were resolved by group discussion. Data were extracted using a predetermined format, including authors, publication year, HRA methods, generalization, domains, and results.

RESULTS

Search Results

Totally 1394 records were retrieved from Web of Science database searching. After 21 duplicates were deleted, 1373 records were left for title and abstract screening, and 1223 were excluded. A total of 150 records were retained for full-text screening against the inclusion and exclusion criteria. In addition, 135 records were removed because of unrelated HRA methods (n = 103), non-English (n = 2), unrelated flight crew tasks (n = 5), analysis of HFCAS only (n = 11), introduction of HRA only (n = 2), and no full text (n = 14). Finally, 13 papers were considered within the scope of this review (Table 1). The process of literature search and selection is shown in Figure 1.

First-Generation HRA Methods in Aviation Safety

This review identified three articles on the first-generation HRA in aviation safety, including THERP (n = 1), HEART (n = 1), and HEART in combination with the Improved Analytics Hierarchy Process (IAHP). Yang et al. (2014) used THERP to analyze flight crew errors during the takeoff phase. In this study, pilot errors were qualitatively analyzed and classified into 10 modes based on hierarchical task analysis (HTA) and interviews with pilots. Based on the event tree and scenario analysis, the pilot errors for each takeoff sub-task were assigned with nominal HEP, and the effects of selected PSFs, including external, internal, and stress factors, on the interdependence among tasks were determined by analysts. The HEP of the whole task was calculated, and fatigue was identified as the major contributing factor.

Kunlun et al. (2011) followed the HEART approach to identify potential human errors associated with deicing operation. The analysis started with HTA based on "AC 91-74B - Pilot Guide: Flight in Icing Conditions" (), followed by assigning tasks to one of nine generic categories regarding working condition of flight crews. Then, the study selected eight of 38 error-producing conditions (EPC) that covered individual (including training, judgements, knowledge, skill, and experience), group (conflicting goals and time), and organization levels (commercial pressure and control systems). Each EPC was

Table 1. HRA study in flight safety (n = 13).

Author (Publication year)	HRA method	Domain	Results
First-generation HRA methods			
Yang, K., et al. (2014)	THERP	Take-off phase of A320 aircraft.	Fatigue was identified as the major contributing factor.
Kunlun, S., et al. (2011)	HEART	Flight crew task, "provide safe flight under icing condition" in FAA AC 91-74A.	The model was validated in a U.S. database with icing and winter weather-related aviation flight incidents in commercial civil flights between 1998 and 2007.
Guo, Y., et al. (2020)	IAHP-HEART	Flight task analysis of B737 aircraft.	The highest HEP was observed in the approach phase followed by the landing phase and the most influential factor was mental strain or loss of attention experienced by pilots.
Second-generation HRA methods			
Hirose, T., et al. (2016)	FRAM-Fuzzy CREAM	American Airline B737 aircraft crash into the terrain near Cali airport in 1995.	The studied accident was caused by a departure from standard operating procedures.
Guo, Y., et al. (2019)	CREAM	Operations in the approach and landing phases of flight task.	The prominent failure types leading to human errors against flight safety were action missed, wrong identification and action of wrong type.
Burns, K., et al. (2020)	SPAR-H	Pilot's action with use of Digital Co-pilot in general aviation.	The model with nominal HEPs for knowledge, rule and skill-based errors caused by pilots after introducing PSF multipliers to explain the corresponding effects of four factors, including "stress", "complexity", "available time", and "ergonomics/human machine interface".
Guo, Y., et al. (2020b)	CREAM	Flight operation process of B737-800 aircraft.	The pilot's performance is reliable during the critical flight tasks, despite a minor deviation.

Continued.

Table 1. Continued.

Author (Publication year)	HRA method	Domain	Results
Third-generation HRA methods			
Bandeira, et al. (2018)	BN	Existing practice in the execution of the tasks performed by the pilot.	Factors in each of the four categories, including "competencies", "performance-shaping factors", "management and organizational factors", and "environmental factors" were identified and related. Relationships among these factors can affect pilot's performance on flight tasks.
Chen, W., et al. (2018)	BN	484 aviation accidents related to flight crew's performance factors occurring from 1999 to 2012 from website data.	"Flying skills", "vigilance", "emergency mishandling", "safety culture", and "crew coordination" were verified as the critical factors for evaluating pilots' performance.
Lyu, T., et al. (2019)	HFACS-BN	142 civil and general aviation accidents /incidents related to ATC human factors worldwide from 1980 to 2019.	The safety performance of Air Traffic Control was the greatest affected by unsafe acts (79.4%), whereas preconditions for safe acts contributed the least influential factor (40.3%).
Guo, X., et al. (2020)	FTA-BN	304 typical flight human error accidents / incidents from 1980 to 2018 from worldwide authorities' reports.	The system has 96 failure modes and the flight human errors in flight were determined by integrating numerous risk factors. The absence of non-technical skills and the deficiencies and deviation of technical skill were found to exert a greater influence on flight human errors. While fundamental events of organizational, environmental, and equipment factors was weakly related to the system.

Continued.

Table 1. Continued.

Author (Publication year)	HRA method	Domain	Results
Meng, B., et al. (2022)	HFACS-BN	74 controlled flight into terrain (CFIT) accidents investigation report from 2001 to 2020.	The precondition for unsafe acts exerted the greatest effect on the controlled flight into terrain accidents. This study also identified that "Inadequate supervision", "intentional noncompliance with standard operating procedures/cross-check", "ground proximity warning system not installed or failure", "adverse meteorological environment", and "ground-based navigation aid malfunction or not being available" were identified as the most contributing factors leading to controlled flight into terrain accidents.
Zhou, Z., et al. (2023)	CATSSR-BN	7,265 incident cases in the Aviation Safety Reporting System (ASRS).	The development of a decision-making model with 10 causal factor events to minimize safety risk in civil aviation.

determined for the Assessed Proportion of Affect values based on research group discussion and for calculating HEP. The model was validated in a U.S. database with icing and winter weather-related flight incidents in civil flights between 1998 and 2007.

Guo and Sun (2020a) combined the HEART and IAHP to improve the prediction accuracy of HEP during aircraft control. By addressing the limitation of HEART based on experts to determine APOA values and weights for each EPC, the IAHP was added to determine the respective influence of subtask-specific EPC via a fuzzy judgment matrix. The new model was tested in selected data between 2000 and 2019 extracted from a national human-related aviation incident database. The results showed that the highest HEP was observed in the approach phase followed by the landing phase and the most influential factor was mental strain or loss of attention experienced by pilots.

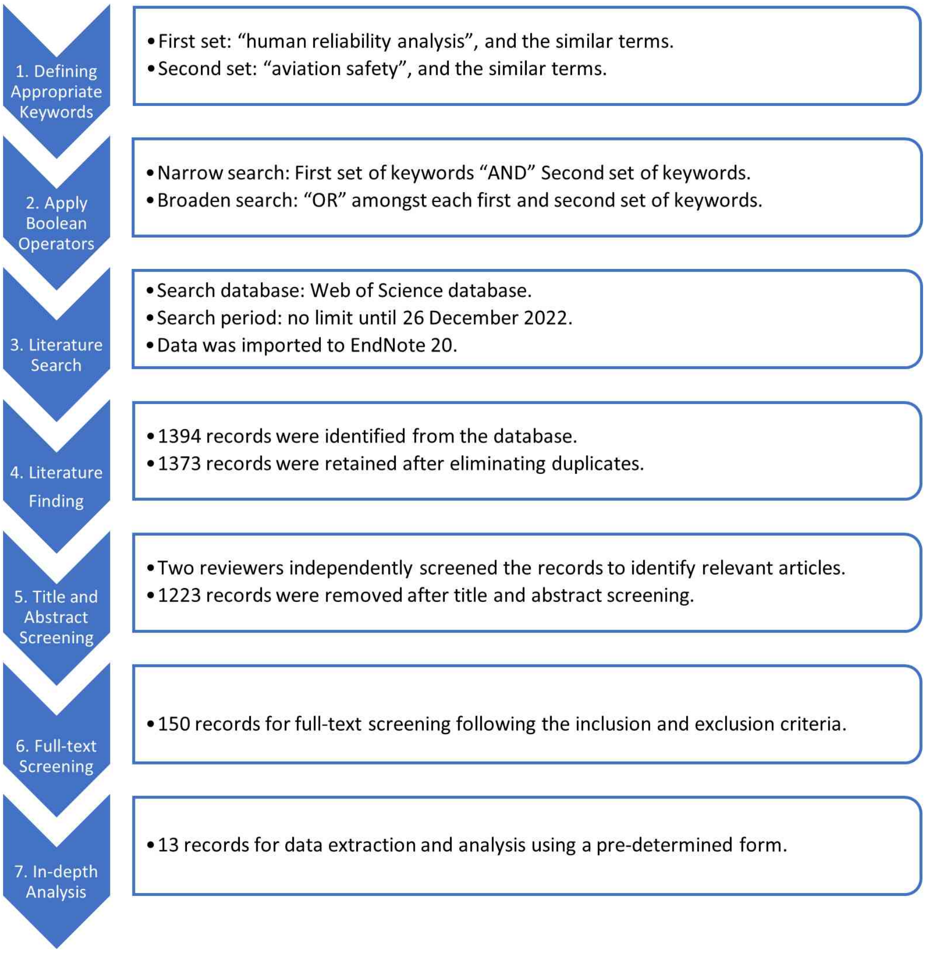

Figure 1: Retrieval process of this study.

Second-Generation HRA Methods in Aviation Safety

Four articles, including SPAR-H (n = 1) and CREAM (n = 3), were identified for the second-generation HRA on aircraft safety. Using the 5-year retrospective data of pilot-related accidents, Burns and Bonaceto (2020) successfully validated the SPAR-H model that estimated nominal HEPs for knowledge-, rule, and skill-based errors caused by pilots after introducing PSF multipliers to explain the corresponding effects of four factors, including "stress", "complexity", "available time", and "ergonomics/human machine interface".

The other three studies were related to the CREAM approach (Hirose, Sawaragi & Horiguchi, 2016; Guo, et al., 2019; Guo & Sun, 2020b). Hirose et al. (2016) used CREAM with the Functional Resonance Analysis Methods (FRAM) to analyze an air crash accident in Colombia in 1995. In this study, five functions of FRAM were defined, and nine Common Performance Conditions (CPC) with two additional factors ("Availability of Resources" and "Quality of Communication" tailored for the study) were added. The weight of each CPC was decided by analysts, and the probability of action failure was calculated on the basis of four control models (strategic, tactical, opportunistic, and scrambled). The study inferred that the accident was caused by a departure from standard operating procedures. Guo et al. (2019) proposed the modified CREAM to analyze four cognitive functions ("observation", "interpretation", "planning", and "execution function") associated with flight tasks particularly in approaching and landing. Nine PIFs were used, with four modified factors, namely, "ground support," "crew workload management," "crew training and experience," and "produce format consistency and verification quality," tailored for the study. The expected affect index, which is the weight of each PIF for the four cognitive functions, was determined by the Delphi method based on the CREAM framework prior to calculating the HEP for operation tasks. The result indicated that the prominent failure types that lead to human errors against flight safety are action missed, wrong identification, and action of wrong type. Compared with HEART, this study with CREAM showed a better fit to the actual situation for assessing human reliability. Guo et al. (2020b) further validated this modified CREAM with real data during the Boeing's flight operation process. The results showed that the pilot's performance is reliable during the critical flight tasks, despite a minor deviation.

Third-Generation HRA Methods in Aviation Safety

Six articles focus on the third-generation method by using the BN-based approaches. Two studies used BN alone, whereas the other four combining BN methodology with HFCAS, Fault Tree Analysis (FTA), and Commercial Air Transportation System Safety Risk (CATSSR) studied pilots' performance for flight procedures (Bandeira, Correia & Martins, 2018; Chen & Huang, 2018; Lyu, Song & Du, 2019; Meng & Lu, 2022; Guo, et al., 2020; Zhou, et al., 2023). Bandeira et al. (2018) used the BN approach to develop a general model for determining key factors leading to aircraft accidents. The results showed that the factors in each of the four categories, including "competencies", "performance-shaping factors", "management and organizational factors", and "environmental factors", were identified and related. Relationships among these factors can affect pilot's performance on flight tasks. The appropriateness of BN was validated in a study of analyzing 484 aviation accidents between 1999 and 2012 (Chen et al., 2018). The results verified "flying skills", "vigilance", "emergency mishandling", "safety culture", and "crew coordination" as critical factors for evaluating pilots' performance.

This review identified two studies with a combination of BN and HFACS. Using data from 142 civil and general flight accidents or incidents worldwide involving Air Traffic Control from 1980 to 2019, Lyu et al. (2019) examined the human factors contributing to Air Traffic Control performance. The results showed that safety performance of Air Traffic Control was the greatest affected by unsafe acts, (79.4%) whereas preconditions for safe acts acting as the least influential factor (40.3%). Meng et al. (2022) studied the causality and intrinsic correlations of 74 controlled flight into terrain accident investigations reported from 2001 to 2020. Results indicated that the precondition for unsafe acts (exerted the greatest effect on the controlled flight into terrain accidents. This study also identified that "Inadequate supervision", "intentional noncompliance with standard operating procedures/cross-check", "ground proximity warning system not installed or failure", "adverse meteorological environment", and "ground-based navigation aid malfunction or not being available" were identified as the most contributing factors leading to controlled flight into terrain accidents.

The other two studies used BN along with other methods, including FTA and CATSSR. Guo et al. (2020) used the FTA-BN approach to evaluate the flight human error related unsafe events incurring in 304 flight accidents or incidents between 1980 and 2018 from worldwide authorities' reports. System risks were identified and analyzed qualitatively by FTA, and the respective probabilities of events of flight human error were identified by BN. The results showed that the system has 96 failure modes, and flight human errors were determined by integrating of numerous risk factors. The absence of nontechnical skills and the deficiencies and deviation of technical skill were found to exert a greater influence on flight human errors. While fundamental events of organizational, environmental, and equipment factors was weakly related to the system. Zhou et al. (2023) used a sample of 7,265 incident cases to test a CATSSR-BN model for evaluating "causal factor events" and the "result events" in particular in-flight transportation system. This resulted in the development of a decision-making model with 10 causal factor events to minimize safety risk in civil aviation.

DISCUSSION

The review found that the first, second, and third generations of HRA were applied in aviation safety. In this review, we found that only one study used THERP, and the other two studies used HEART. Despite being the pioneer HRA method, THERP was largely subjective, looking at selected PSFs for a particular task. The HEART has been empirically validated in existing databases, indicating the potential for flight safety risk prevention and reduction.

Although HEART can successfully assess human-related flight operation errors, it still has limitations. The HEART does not fully consider individual cognitive errors from actual flight task situations. For example, the disturbance between pilots and the flight deck interface is not included; thus, pilots' cognitive performance is affected. Furthermore, this technique is limited in uncovering reasons for human errors regarding the flight sub-task operation but not the judgement or decision-making.

The review included four articles for the second-generation HRA, one study for SPAR-H and three studies for CREAM. The SPAR-H study was validated with accident rate data, but it was only derived from the results of pilots' use of a digital co-pilot system. Furthermore, this study used retrospective analysis and future prospective investigation of human reliability, and flight safety improvements are necessary. In addition, the review shows that the modified CREAM is more sensitive to detecting individual cognitive errors than HEART for critical flight tasks. However, only one article had such validation data, indicating that this method currently lacks empirical validation. Although the second-generation methods are built on cognitive modelling enabling an in-depth explanation of human errors, they have serious drawbacks (Kim et al., 2006). For example, the models lack of empirical testing and validation, leading to questioning the validity issues. The selection of PSF is highly determined by expert consensus, which is qualitative in nature. By addressing these limitations of second-generation HRA, third-generation methods have achieved a progress moving towards real data-based testing using advanced methodological tools to improve the accuracy of HRA model (Di et al., 2013). However, in the present review, only BN-based individual or combined approaches were applied in the literature. Notably, the other third-generation HRA methods such as PHOENIX and IDHEA have not been used.

CONCLUSION

By providing a picture of HRA application in aviation accidents, this review showed that the third-generation HRA with BN was frequently reported in the literature. Testing the other third-generation HRA models driven by data is a future direction in the field of airworthiness.

FUNDING

This work was supported by the National Natural Science Foundation of China [grant number 72071170].

ABBREVIATIONS

APOA	Assessed proportion of affect values	HEF	Human failure events
BN	Bayesian network	HEP	Human error probability
CPA	Cognitive performance analysis	HTA	Hierarchical task analysis
CPC	Common performance condition	IAPH	Improved analytics hierarchy process
CREAM	Cognitive reliability and error analysis method	MMI	Man–machine interface
EPC	Error-producing conditions	PAF	Probability of action failure
FRAM	Functional resonance analysis methods	PAR	Probabilistic risk assessment
FTA	Fault Tree Analysis	PIF	Performance influencing factor
HEART	Human error assessment and reduction technique	PII	Performance impact index
HFACS	Human factor analysis and classification system	SPAR-H	Standardized plant analysis risk human reliability analysis
HRA	Human reliability analysis	THERP	Technique for human error rate prediction

REFERENCES

Abeyratne, R., & Abeyratne, R. (2012). Aircraft accident and incident investigation. *Air Navigation Law*, 105–111.

Bandeira, M. C. G. S. P., Correia, A. R., & Martins, M. R. (2018). General model analysis of aeronautical accidents involving human and organizational factors. *Journal of Air Transport Management*, *69*, 137–146.

Boyd, D. D. (2017). A review of general aviation safety (1984–2017). *Aerospace Medicine and Human Performance*, 88(7), 657–664.

Burns, K., & Bonaceto, C. (2020). An empirically benchmarked human reliability analysis of general aviation. *Reliability Engineering & System Safety*, *194*, 106227.

Chen, W., & Huang, S. (2018). Evaluating flight crew performance by a bayesian network model. *Entropy*, *20*(3), 178.

Erjavac, A. J., Iammartino, R., & Fossaceca, J. M. (2018). Evaluation of preconditions affecting symptomatic human error in general aviation and air carrier aviation accidents. *Reliability Engineering & System Safety,* 178, 156–163.

Guo, X., Meng, B., Liu, Y., Lu, N., Fu, S., Si, Q. (2020). A human error mechanism for pilot based on fault tree analysis and Bayesian network. *Journal of Intelligent & Fuzzy Systems*, *38*(6), 6863–6871.

Guo, Y., & Sun, Y. (2020a). Flight safety assessment based on an integrated human reliability quantification approach. *Plos one*, *15*(4), e0231391.

Guo, Y., & Sun, Y. (2020b). Human reliability quantification in flight through a simplified CREAM method. In *Advances in Human Factors of Transportation: Proceedings of the AHFE 2019 International Conference on Human Factors in Transportation, July 24-28, 2019, Washington DC, USA 10* (pp. 762–773). Springer International Publishing.

Guo, Y., Sun, Y., Yang, X., & Wang, Z. (2019). Flight safety assessment based on a modified human reliability quantification method. *International Journal of Aerospace Engineering*, *2019*.

Hirose, T., Sawaragi, T., & Horiguchi, Y. (2016). Safety analysis of aviation flight-deck procedures using systemic accident model. *IFAC-PapersOnLine*, *49*(19), 19–24.

Hou, L. X., Liu, R., Liu, H. C., & Jiang, S. (2021). Two decades on human reliability analysis: a bibliometric analysis and literature review. *Annals of Nuclear Energy*, *151*, 107969.

Kunlun, S., Yan, L., & Ming, X. (2011). A safety approach to predict human error in critical flight tasks. *Procedia Engineering*, *17*, 52–62.

Lyu, T., Song, W., & Du, K. (2019). Human factors analysis of air traffic safety based on HFACS-BN model. *Applied Sciences*, *9*(23), 5049.

Meng, B., & Lu, N. (2022). A hybrid model integrating HFACS and BN for analyzing human factors in CFIT accidents. *Aerospace*, *9*(11), 711.

Mkrtchyan, L., Podofillini, L., & Dang, V. N. (2015). Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability engineering & system safety, 139,* 1–16.

Pan, X., Lin, Y., & He, C. (2017). A review of cognitive models in human reliability analysis. *Quality and Reliability Engineering International, 33*(7), 1299–1316.

Reason, J. (2016). Managing the risks of organizational accidents. Routledge.

Tao, J., Qiu, D., Yang, F., & Duan, Z. (2020). A bibliometric analysis of human reliability research. *Journal of Cleaner Production*, *260*, 121041.

Transportation Safety Board of Canada (2023). Statistical summary: Air transportation occurrences in 2016. Website: https://www.tsb.gc.ca/eng/stats/aviation/2016/ssea-ssao-2016.html

Wan M, Liang Y, Yan L, Zhou T. (2021). Bibliometric analysis of human factors in aviation accident using MKD. *IET Image Processing, 2021,* 1–9.

Wiegmann, D. A., & Shappell, S. A. (2003). A human error approach to aviation accident analysis: The human factors analysis and classification system. Routledge.

Yang, K., Tao, L., & Bai, J. (2014). Assessment of flight crew errors based on THERP. *Procedia engineering*, *80*, 49–58.

Zarei, E., Yazdi, M., Abbassi, R., & Khan, F. (2019). A hybrid model for human factor analysis in process accidents: FBN-HFACS. *Journal of loss prevention in the process industries*, *57*, 142–155.

Zhou, T., Zhang, J., & Baasansuren, D. (2018). A hybrid HFACS-BN model for analysis of Mongolian aviation professionals' awareness of human factors related to aviation safety. *Sustainability, 10*(12), 4522.

Zhou, Z., Yu, X., Zhu, Z., Zhou, D., & Qi, H. (2023). Development and application of a Bayesian network-based model for systematically reducing safety risks in the commercial air transportation system. Safety science, 157, 105942.