

Research Article

Flight Safety Assessment Based on a Modified Human Reliability Quantification Method

Yundong Guo , Youchao Sun , Xiufang Yang, and Zongpeng Wang

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Correspondence should be addressed to Youchao Sun; sunyc@nuaa.edu.cn

Received 18 April 2019; Revised 14 August 2019; Accepted 6 September 2019; Published 27 October 2019

Academic Editor: Jeremy Straub

Copyright © 2019 Yundong Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human error is one of the most important risk factors affecting aviation safety. The original Cognitive Reliability and Error Analysis Method (CREAM) developed for the nuclear industry is reliable for human reliability quantification, but it is not fully applicable to human reliability analysis in aviation because it neglects the characteristics of long-duration flights. Here, we propose a modified CREAM method to predict human error probability in flight and provide some improvement measures for critical operations. A set of performance influencing factors (PIFs), such as flight procedures and ground support, is established to reflect operational scenarios in flight. Then, we develop the expected affect index of PIFs and the Scenario Influence Index to construct a quantitative model of human reliability. The probability of human error for each operation in the approach and landing phases is obtained with the modified CREAM method, and the results indicate that the most important cognitive function that influences human reliability is missed action. The proposed method may be a suitable tool for human reliability quantification in aviation considering long-duration flights. The method also has great practical significance for improving flight safety.

1. Introduction

With the continuous emergence of advanced avionics technology and the increase in the reliability of onboard equipment, the aviation accident rate related to the electro-mechanical systems of aircraft has sharply decreased in recent decades [1, 2]. However, the aviation accident rate associated with human errors has not yet been well controlled. The accident statistics reported by the International Civil Aviation Organization (ICAO) show that over 70% of accidents are directly or indirectly related to human errors [3]. The effective assessment of human reliability in complex tasks is essential in assisting analysts in identifying critical human errors and improving human reliability.

The role of pilots has become managing in pilot-aircraft interactions instead of executing tasks, as the cockpit interface has become largely automatic, integrated, and intelligent. The pilot requires more diagnostics than just routine operations in flight, so human cognitive performance may have an important impact on flight safety [4]. The Cognitive Reliability and Error Analysis Method (CREAM) was first developed by Hollnagel for human reliability analysis in the

nuclear power industry [5]. This method can be utilized to proactively predict potential human error and retroactively quantify human error. CREAM mainly uses the Contextual Control Model (COCOM), which is used to select actions and assumes that the level of control of the operator is variable to determine the probability of human error. COCOM includes four control modes, including scrambled control, opportunistic control, tactical control, and strategic control. The higher the control level of the operator, the higher the operator's performance reliability. CREAM uses a classification scheme consisting of a number of groups reflecting the causes of erroneous actions. The classification scheme can be utilized to anticipate and describe how potential human errors occur. Furthermore, this method defines the relation between the consequences of the error and the causes. Thus, CREAM uses a systematic approach to identify and quantify human error. However, the classification scheme describing the causes of erroneous actions in the CREAM framework is mainly applicable for nuclear power plants and is difficult to directly apply to other domains. Moreover, CREAM does not offer reasonable recommendations for reducing the occurrence of errors and is complicated to apply.

Based on the CREAM approach, numerous modified methods have been proposed and applied. Yang et al. [6] proposed a modified CREAM method using a fuzzy Bayesian approach to quantify the reliability of marine engineers. Wang et al. [7] developed a clonal selection algorithm to evaluate linguistic variables in CREAM, and it was successfully applied in a safety assessment of power systems. Calhoun et al. [8, 9] introduced new performance shaping factors to quantify human error probability in space missions. Chen et al. [10] used CREAM and a Bayesian Network to evaluate human unreliability in space missions. Wu et al. [11] proposed an evidential reasoning-based CREAM method to assess human reliability in a ship capsizing accident. Bedford et al. [12] introduced PSF weights with nominal probabilities to improve the sensitivity and uncertainty of the original CREAM method. A weighted CREAM model based on fuzzy logic theory was developed to enhance the logicity among PSFs and cognitive functions in marine accidents [13]. A THERP-CREAM and expert opinion auditing method that effectively reflects plant reality and fills data gaps was presented in [14]. The Fuzzy Analytic Hierarchy Process (FAHP) was incorporated into CREAM to assess seafarer reliability in tanker shipping [15]. Numerous human reliability assessment techniques have proven to be reliable in terms of applications involving power plants, space missions, and marine transportation. However, the application of such methods for human error quantification in flight safety assessment has not yet been documented in journals.

Considering the operational characteristics in long-duration flights, such as flight procedures and ground support, the aforementioned methods are difficult to directly apply to analyze and quantify human error in flight. We propose a modified CREAM approach to assess human reliability in flight and provide some improvement measures for critical operations. In comparison with the existing studies, our model includes an applicable classification scheme consisting of a number of groups that describe the causes of errors in aviation. This method is easy to use in practical applications. Moreover, improvement measures for operations with high error probability are proposed to enhance flight safety.

2. Methodology

2.1. Performance Influencing Factors. The original CREAM method defines a classification scheme consisting of nine groups that describe the causes of erroneous actions [16]. These erroneous causes, which are known as performance influencing factors (PIFs), involve the adequacy of organization, adequacy of the man-machine interface, working conditions, availability of procedures, number of simultaneous goals, available time, time of day, and crew collaboration quality. These factors are mainly used for probabilistic safety assessments in nuclear power plants. However, flight scenarios, particularly the scenario of long-duration flight, are clearly different from scenarios in the nuclear industry. During a long-duration flight, each pilot must be familiar with various flight information, such as the aircraft status, fault reservation information, alternate airports, airport weather,

and air routes. Numerous and complex operations should also be performed exactly and at the appropriate times in each phase of a long-duration flight. In the meantime, the crew must communicate regularly with support teams on the ground via wireless communication to ensure flight safety. These tedious operations pose a great challenge to human physiology and psychology during flight, which may further affect human performance and even lead to human error. Therefore, it is necessary to define a set of appropriate PIFs that describe the causes of erroneous actions in flight and assess human reliability.

According to the characteristics of flight, there are four important factors that affect human performance during a long-duration flight: ground support, crew workload management, crew training/experience, and procedure format consistency and verification quality. The modified “ground support” and “crew workload management” PIFs are different from the “crew collaboration quality” factor considered in the original method, which represent the level of collaboration between crew members and whether the crew executes well under pressure. The “ground support” PIF mitigates risk and improves crew collaboration quality. Members of a ground support team can act as extra sets of eyes and ears, not only issuing effective commands but also monitoring and verifying data to reduce crew workload and fatigue [17]. Meanwhile, the crew can aid in malfunction procedures and even handle situations unpredicted by ground support. The “crew workload management” PIF involves the crew’s consideration of conflicts over downlink priority under various conditions, crew flight time constraints, and other risk mitigation requirements. It concerns how the crew optimizes resource usage in different flight phases. The effective management of crew time and workload can reduce human error and improve flight efficiency. In addition, the “crew workload management” PIF reflects human performance in terms of human cognitive activities. Therefore, the “ground support” and “crew workload management” PIFs may be more suitable for capturing the task characteristics of a long-duration flight than the “crew collaboration quality” PIF is. The definition of the “adequacy of training and preparation” PIF in the original CREAM framework states that insufficient skills or knowledge can be a cause of human error [5]. However, even highly trained crews cannot account for all potential situations. The modified “crew training and experience” PIF concerns the applicability, recency, and repetition of training as well as the crew’s prior experience. The more applicable or closer in time the flight task is to training, the lower the risk. Repetition of training can also decrease the flight risk. The prior experience of the crew comprises official training, simulator training, and flight experience. In addition, there will inevitably be some differences between training procedures and actual missions, such as the presence of a microgravity environment and spatial disorientation. Although the possession of sufficient skills or knowledge is important for flight safety, rich experience is also an essential factor. The definition of the “availability of procedures” PIF, proposed by Hollnagel, states that procedural deficiencies or discrepancies are an important factor in erroneous human actions [5]. Based on this PIF in the original CREAM framework,

we develop a modified PIF called “procedure format consistency and verification quality.” This factor includes not only the consistency of procedure formats but also the procedure verification quality. Procedures may be jointly developed by different organizations or airlines, but their formats are likely to be inconsistent due to lack of aviation training or differing operation standards. Thus, it is necessary to consider the consistency of procedure formats. High-quality procedures can reduce the risk of human error, but it is difficult to verify all possible procedures on the ground because of resource constraints. Therefore, the proposed “procedure format consistency and verification quality” PIF is different from the original “availability of procedures” PIF.

Thus, we optimize the PIFs in the original CREAM framework to propose 9 PIFs for long-duration flight (Table 1). The 9 proposed PIFs can be classified into four types describing different aspects of an actual situation: manpower, machine, environment, and task factors. The manpower factors include ground support, crew workload management, crew training and experience, and adequacy of organization. These factors reflect the cognitive competence and operational capability of the crew as well as the communication and collaboration between the crew and the ground supporters. The machine factor is the adequacy of the man-machine interface, which is a critical channel in man-machine interaction. It is a key cause of erroneous human activities. The work conditions factor is an environment factor and thus directly affects human physiological and psychological activities as well as operating comfort. A specific task must be performed in a closed man-machine-environment loop. The task factors, including procedure quality/quantity and the time stress for task completion, also have a significant influence on human performance. It is impossible for the proposed PIFs to cover all potential error factors, but these 9 modified PIFs can reasonably reflect the majority of the factors that influence human performance in terms of man-machine-environment interaction. Each of the proposed PIFs has different levels. For example, the levels of ground support are very efficient, efficient, inefficient, and deficient. The influence of a PIF on performance can be classified as improved, not significant, or reduced performance.

The human information process mainly involves four cognitive functions, including observation, interpretation, planning, and execution functions [18]. Each PIF has a different weight for each of these four functions. For example, if the ground support is deficient, the failure probabilities of the observation and interpretation functions will not change, but those of the plan and execution functions may increase. In theoretical studies and engineering practice, it is often necessary to establish models and ignore the corresponding complexity [19]. One PIF may have different influences on these four functions, and the weight of each cognitive function is not easy to clearly define due to the high complexity and uncertainty of the information process [12, 20]. Therefore, the weight factors of the different levels of the PIFs for the four functions can be simplified to a single number [21], which is defined as the Expected Affect Index (EAI). The values of the EAI are mainly based on the weight factors provided by the original CREAM framework and further

determined by the Delphi Method. In brief, we selected 10 experts to score the PIF weights in a range of -3 to 3, or from extremely important to extremely unimportant. The survey was conducted in three rounds, and the average expert score was used as the weight of the PIF. Table 2 lists the final EAI values for the nine PIFs. It is noteworthy that these values are used as a reference and flexible rather than constant in practical applications.

2.2. Human Reliability Quantitative Model

2.2.1. The Basic CREAM Method.

There are four control modes that reflect the human reliability status in the original CREAM method: strategic, tactical, opportunistic, and scrambled modes. The strategic control mode is defined as the operator having sufficient time to take the entire situation into account and plan accordingly. The operator performs task plans in accordance with standard procedures or rules in tactical control mode. The opportunistic control mode refers to human performance being driven by interface characteristics or the most commonly used action of the operator, which is also known as gambling heuristics. The scrambled control mode refers to the operator being unfamiliar with the current situation and losing situational awareness when the task demands are very high. Moreover, each control mode has a corresponding human error probability interval. For example, the interval of the human error probability for strategic control is from $5.0E-5$ to $1.0E-2$; the probability interval for tactical control is from $1.0E-3$ to $1.0E-1$; the probability interval for opportunistic control is from $1.0E-2$ to $5.0E-1$; and the probability interval for scrambled control is from $1.0E-1$ to $1.0E-0$ [5].

In the basic CREAM method, the control mode is determined based on an array of combined PIFs expressed as $(\sum_{\text{reduced}}, \sum_{\text{improved}})$ [5]. \sum_{reduced} is the number of PIFs that reduce performance reliability, and \sum_{improved} is the number of PIFs that improve performance reliability. Figure 1 depicts the relation between the PIF and control mode. For example, (1, 6) implies that the control mode of the operator is strategic mode. In this paper, we introduce the Scenario Influence Index (SII), denoted by α , to describe the influence of the PIFs on human reliability. The SII can be expressed as follows.

$$\alpha = x - y = \sum_{\text{reduced}} - \sum_{\text{improved}} \quad (1)$$

First, Equation (1) is transformed into the linear function $y = x - \alpha$, where $x \in [0, 9]$, $y \in [0, 7]$. Then, the inclined lines are plotted, as shown in Figure 1. The α value generally represents the control mode with only three exceptions. However, due to the potential overlap of two adjacent control modes, it is difficult to determine which mode is absolutely correct among the continuous changes. Therefore, it is reasonable and acceptable to use α to determine the control mode, even with three exceptions. Consequently, the corresponding relation between α and the control mode is illustrated in Table 3. These data can be used to determine the control mode of the operators in the basic CREAM method.

TABLE 1: Modified PIFs for flight.

PIF	Description
Ground support*	The level of support and resources provided by the ground controllers for the task or work being performed. Very efficient/efficient/inefficient/deficient Improved/not significant/reduced
Crew workload management*	The level of crew workload to maximize efficacy. Supportive/adequate/tolerable/inappropriate Improved/not significant/reduced
Crew training and experience*	This factor involves the applicability, recency, and repetition of training as well as the crew's prior experience. High experience/limited experience/inadequate Improved/not significant/reduced
Procedure format consistency and verification quality*	This factor mainly refers to the consistency and effectiveness of the procedure formats as well as how the ground verification quality influences human activities. Appropriate/acceptable/inappropriate Improved/not significant/reduced
Procedure quantity	The number of simultaneous goals that need to be completed in a given period. Fewer than capacity/matching current capacity/more than capacity Not significant/not significant/reduced
Time stress	The time allowed by the system and the actual time required to complete the task. Adequate/temporarily inadequate/continuously inadequate Improved/not significant/reduced
Adequacy of the man-machine interface	The information availability and identifiability in the display interface. Supportive/adequate/tolerable/inappropriate Improved/not significant/reduced
Adequacy of organization	The quality level of organizational factors, such as policies, planning, management, and quality assurance, that affect individual safety consciousness. Very efficient/efficient/inefficient/deficient Improved/not significant/reduced
Working conditions	This factor generally refers to the physical working environment, including glare, large vibrations, alarms, uncomfortable temperatures, and external interference events. Advantageous/compatible/incompatible Improved/not significant/reduced

* indicates a cause of erroneous actions in flight.

When the number of reduced PIFs is equal to the number of improved PIFs, α is equal to zero. The human error probability (HEP) in this situation is considered to be the basic HEP, which is denoted as HEP_0 . It is assumed that the relationship between the crew error probability and external environment is a logarithmic function [21] expressed as follows.

$$\log \left(\frac{HEP}{HEP_0} \right) = n\alpha. \quad (2)$$

In this case, n is a constant coefficient that can be calculated as follows:

$$\begin{cases} \log \left(\frac{HEP_{\max}}{HEP_0} \right) = n\alpha_{\max}, \\ \log \left(\frac{HEP_{\min}}{HEP_0} \right) = n\alpha_{\min}. \end{cases} \quad (3)$$

Equation (3) can be rearranged to obtain

$$n = \log \frac{HEP_{\max}/HEP_{\min}}{\alpha_{\max} - \alpha_{\min}}. \quad (4)$$

The maximum value of reduced PIFs is 9, and the maximum value of improved PIFs is 7. Therefore, we can obtain $\alpha_{\max} = 9$, $\alpha_{\min} = -7$. If $(HEP_{\max} = 1, HEP_{\min} = 5.0E - 4)$ based on the original CREAM database, then $n = 0 : 21$ and $HEP_0 = 0.0129$. Finally, Equation (2) can be represented as follows.

$$HEP = HEP_0 \times 10^{0.21\alpha} = 0.0129 \times 10^{0.21\alpha}. \quad (5)$$

In the basic CREAM method, Equation (2) indicates that each PIF is equally important for human performance, and this equation can only be used in the initial screening phase. The quantification of specific task operations cannot be performed because the difference in the impact of each PIF on human reliability is not considered.

TABLE 2: EAI values of PIFs.

PIF	Level	Modified EAI
Ground support	Very efficient	-1.8
	Efficient	0
	Inefficient	1.0
	Deficient	1.8
Crew workload management	Supportive	-1.8
	Adequate	-1.0
	Tolerable	0
	Inappropriate	1.8
Crew training and experience	High experience	-1.8
	Limited experience	0
	Inadequate	1.8
Procedure format consistency and verification quality	Appropriate	-1.6
	Acceptable	0
	Inappropriate	1.6
Procedure quantity	Fewer than capacity	0
	Matching the current capacity	0
	More than capacity	1.2
Time stress	Adequate	-1.4
	Temporarily inadequate	1.0
	Continuously inadequate	2.5
Adequacy of the man-machine interface	Supportive	-1.4
	Adequate	-0.6
	Tolerable	0
	Inappropriate	1.8
Adequacy of organization	Very efficient	-0.6
	Efficient	0
	Inefficient	0.6
	Deficient	1.0
Working conditions	Advantageous	-0.8
	Compatible	0
	Incompatible	1.6

2.2.2. *The Extended CREAM Method.* In the human cognitive information-processing model, there are four phases of man-machine interaction, including the observation, interpretation, planning, and execution phases. The effect of PIFs on the four cognitive functions may vary, but the differences depend heavily on expert judgment. Furthermore, the human error mode is not easy to predict due to the complexity and uncertainty of operations. The prediction range mainly relies on comprehensive analyses of the previous tasks. To simplify the quantification of HEP for each task operation, the functional relationship between the EAI and the SII is developed as follows:

$$\alpha = \sum_{i=1}^9 \theta_i, \quad (6)$$

where θ_i is the EAI value of the i th PIF and its value is presented in Table 2. In the extended CREAM method, Equation (5) is converted to as follows:

$$\text{HEP} = \text{CFP}_0 \times 10^{0.21 \sum_{i=1}^9 \theta_i}, \quad (7)$$

where CFP_0 is the nominal cognitive failure probability, which is defined as a nominal value. Table 4 lists the nominal cognitive error probability data from the original CREAM method [5].

Both the extended method and the basic method assume that tasks occur in a specific scenario, and human reliability is greatly influenced by the task scenario. However, the basic method directly uses the basic HEP to determine the human error mode, and the extended method utilizes the nominal

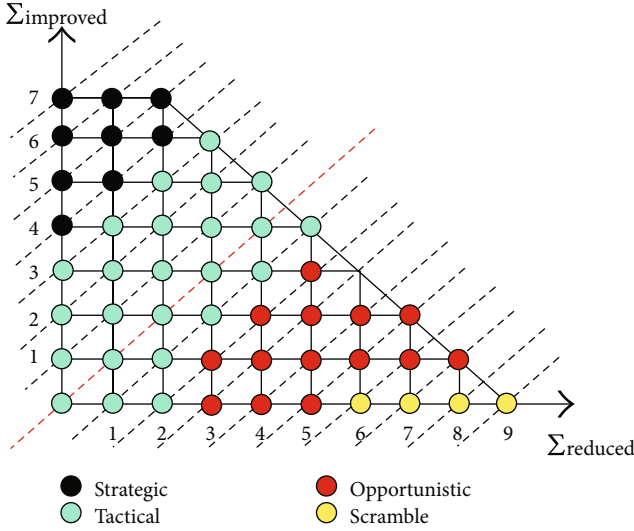


FIGURE 1: Control modes.

TABLE 3: Relation between the SII and control mode.

Control mode	SII
Strategic	(-7, -4)
Tactical	(-3, 1)
Opportunistic	(2, 5)
Scrambled	(6, 9)

TABLE 4: Nominal cognitive error probability.

Cognitive function	Generic failure type	Nominal value
Observation	O1. Wrong object observed	$1.0E-3$
	O2. Wrong identification	$7.0E-2$
	O3. Observation not made	$7.0E-2$
Interpretation	I1. Faulty diagnosis	$2.0E-1$
	I2. Decision error	$1.0E-2$
	I3. Delayed interpretation	$1.0E-2$
Planning	P1. Priority error	$1.0E-2$
	P2. Inadequate plan	$1.0E-2$
Execution	E1. Action of wrong type	$3.0E-3$
	E2. Action at wrong time	$3.0E-3$
	E3. Action on wrong object	$5.0E-4$
	E4. Action out of sequence	$3.0E-3$
	E5. Missed action	$3.0E-2$

value of cognitive function failure to calculate the HEP of each operation.

2.2.3. *Calculation of the HEP for Tasks.* Task analysis is the foundation of the CREAM method. First, the overall task goals are decomposed into a series of operations to support reliability analysis. Then, based on the operation sequence,

TABLE 5: Calculation principles for the HEP of a task.

Logic relation between operations	Dependence between operations	HEP of a task
Parallel operations	Dependence	$HEP_{task} = \min HEP_{operation}$
	Independence	$HEP_{task} = \prod HEP_{operation}$
Serial operations	Dependence	$HEP_{task} = \max HEP_{operation}$
	Independence	$HEP_{task} = \sum HEP_{operation}$

the HEP of each operation is determined with Equation (7). Finally, some principles for combining the HEPs of operations are provided to calculate the HEP of the overall task, as listed in Table 5.

If the operation sequence in a task is parallel and the relation between operations is dependent, the HEP of the overall task is defined as the minimum HEP of all operations. For operations that are independent and/or parallel, the overall HEP of a task is the product of all the operation error probabilities. Similarly, for serial operations with dependence, the maximum HEP of all operations is regarded as the task HEP. When the relation between the serial operations is independent, the total task HEP is assigned as the sum of the HEP values of all operations.

3. Case Study

3.1. *Flight Task Analysis.* To demonstrate the application of the modified method in flight tasks, an appropriate event tree should be constructed to provide support for further human error risk assessment. The flight task event tree is presented in Figure 2 according to the analysis of standard flight procedures and processes. The failure of a task is denoted as F, and the success of a task is denoted as S. Each task has two states: failure and success. If the taxi task fails, the entire flight task fails. After the taxi task is successful, the takeoff task is performed. Similarly, only when the previous task succeeds can the next task be executed.

In the structure of the event tree, there are seven function top events, including taxi, takeoff, stabilized climb, climb and cruise, descent, approach, and landing events. In this section, the task events of approach and landing are selected to illustrate the process of the proposed method. The reasons for choosing the two events are threefold. (1) The accident rate during these two phases is very high in all flight tasks. According to the statistics reported by the Boeing Commercial Airplanes, the percentage of fatal accidents and onboard fatalities in these two flight phases was over 50% from 2008 to 2017 [22]. (2) The environment is uncertain and complex. The approach and landing processes often involve low-visibility weather with clouds, fog, and rain. Furthermore, low-level wind shear is also a common occurrence. All these factors can greatly influence the crew performance reliability. (3) The flight crew plays an important role in these phases. The crew are required to perform more operations correctly than are required in other flight phases. Consequently, the

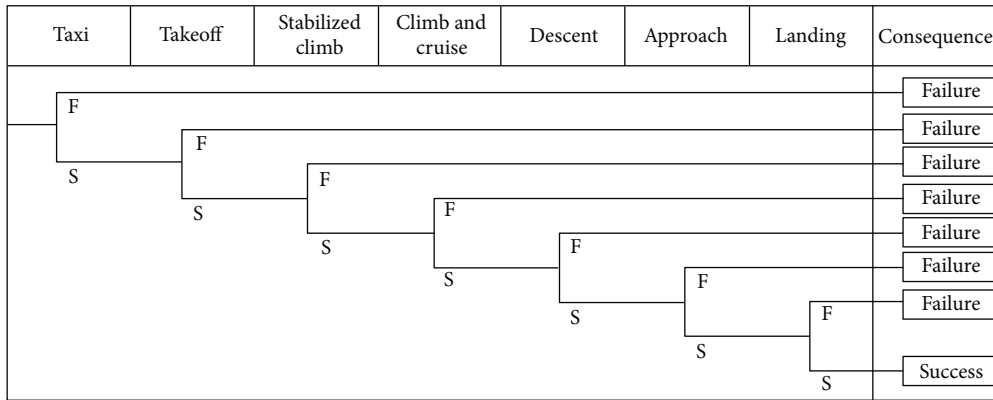


FIGURE 2: Flight task event tree.

TABLE 6: Operations in the approach and landing phases.

Operation description
1. Approach
1.1 Set passenger signs as needed
1.2 Set LANDING light switches to ON at or above 10,000 ft
1.3 Set and crosscheck altimeters at the transition level
1.4 Update arrival and approach procedures as needed
1.5 Update required navigation performance as needed
1.6 Update approach briefing as needed
1.7 Call and do the "APPROACH CHECKLIST"
2. Landing
2.1 Verify cabin is secure
2.2 Call and set flap lever
2.3 Monitor flap and slat extension
2.4 Verify the instrument landing system and identify when on the localizer intercept heading
2.5 Verify LOCALIZER and GLIDE SLOPE are shown when on the localizer intercept heading
2.6 Use HDG SEL to intercept final approach course as needed
2.7 Verify that localizer is captured
2.8 Call "GLIDE SLOPE ALIVE"
2.9 Call "GEAR DOWN" and "FLAP 15"
2.10 Set landing gear lever to DN
2.11 Verify that green landing gear indicator lights are illuminated
2.12 Set flap lever to 15
2.13 Set engine start switches to CONT
2.14 Set speed brake lever to ARM
2.15 Verify that SPEED BRAKE ARMED light is illuminated
2.16 Call and set flap lever at glide slope capture
2.17 Set missed approach altitude on the MCP
2.18 Call and do "LANDING CHECKLIST"
2.19 Verify the crossing altitude at final approach
2.20 Monitor approach
2.21 Verify callouts and autoland status at 500 ft

detailed operations in the approach and landing phases based on the task analysis are listed in Table 6.

3.2. Calculation of the HEP Based on the Basic CREAM Method. Once the flight task analysis has been thoroughly completed, the expected effect of each task scenario on human reliability, known as the PIF level, can be qualitatively

determined based on expert experience and judgment. It is assumed that the task scenario does not change between the approach and landing stages because these two tasks are performed within a relatively short time [23]. Thus, the qualitative results can be obtained as shown in Table 7.

According to the predetermined flight path, air traffic controllers working on the ground act as the eyes and ears of aircraft to command the flight crew to fly correctly. Furthermore, they can provide safety guidance for the flight crew in the entire flight task. Thus, the PIF level of ground support is considered to be "Efficient" under normal conditions. Considering human performance and flight safety, a passenger airliner is generally equipped with at least two flight crew members. Two pilots are required to coordinate and complete the flight mission, and the allocation of crew tasks is reasonable. Thus, the PIF level of crew workload management is deemed to be "Adequate." The flight crew is considered skilled and experienced, as reflected in previous performance reviews. The approach and landing tasks are also regularly performed in training with a flight simulator. Thus, the PIF level of crew training and experience is regarded as "High experience." Standard flight procedures have been successfully used in passenger aircraft for many years. Therefore, the procedure qualities for approach and landing are considered "Appropriate." Although there are hundreds of operating procedures throughout a flight, the crew can still complete these operations smoothly with the assistance of automatic equipment and air traffic controllers. Therefore, the procedure quantity is regarded as "Matching the current capacity."

The time of approach and landing is only approximately 7 minutes of the entire flight process, but the crew is required to correctly perform urgent and complex operations during this limited period. Pilots should diagnose the event and take active measures to ensure the flight safety of approach and landing in less than 7 minutes. The available time is compared with the nominal time to determine the time stress. According to experimental research with a flight simulator, the nominal time of approach is 4 minutes, and the nominal time of landing is 1 minute. A great deal of experience has shown that the simulator-based experiments are ideal and that the results should be adjusted by a factor larger than 1

TABLE 7: PIF levels in the approach and landing stages.

PIF	PIF level in approach stage	PIF level in landing stage
Ground support	Efficient	Efficient
Crew workload management	Adequate	Adequate
Crew training and experience	High experience	High experience
Procedure format consistency and verification quality	Appropriate	Appropriate
Procedure quantity	Matching current capacity	Matching current capacity
Time stress	Temporarily inadequate	Temporarily inadequate
Adequacy of the man-machine interface	Adequate	Adequate
Adequacy of organization	Efficient	Efficient
Working conditions	Incompatible	Incompatible

[24]. Assuming this factor is 1.5, the time of approach and landing will be greater than 7 minutes. Thus, the PIF level of the time stress is deemed “Temporarily inadequate.”

The principles of human-centered cockpit design ensure good visibility and accessibility for pilots. The man-machine interface is simple, clear, and intuitive. Additionally, the control panel of each device is arranged in a compact layout in the cockpit. Thus, the man-machine interface is considered “Adequate.” Airlines are operating under strict conditions, so the organization factor is deemed “Efficient.” There may be severe environmental conditions, such as wind shear or low-visibility weather, during the approach and landing phases, and such conditions can seriously affect the performance reliability of the pilot. Thus, the PIF level of working conditions is regarded as “Incompatible” in such a case.

After all the PIF levels are determined, the SII value is calculated with Equation (1) based on the basic CREAM method. We can obtain $\alpha = 0$. Then, the control modes of the approach and landing tasks are both obtained. Finally, the HEP of the two tasks is calculated with Equation (5), and the results are as follows:

$$\text{HEP} = \text{CFP}_0 \times 10^{0.21\alpha} = 0.0129 \times 10^{0.21 \times 0} = 0.0129. \quad (8)$$

3.3. Calculation of the HEP Based on the Extended CREAM Method. The extended CREAM method is used to further quantify the HEP of each operation. In this section, the weight factor for each PIF level, known as the EAI value, is determined in accordance with the qualitative description provided in the basic CREAM framework. Then, the EAI of each PIF for the two flight tasks is assigned in order based on the qualitative results in Table 7 and the EAI values for the PIFs in Table 2. The final EAI results are obtained as shown in Table 8.

After all the EAI values of the PIFs are determined, the cognitive activity associated with each operation needs to be selected based on the general definitions of cognitive activities provided in the original CREAM framework. These cognitive activities mainly include the following: coordinate, communicate, diagnose, execute, monitor, and verify [5]. Then, the possible cognitive failure type for each cognitive activity can be identified based on Table 4. The results are listed in Table 9 and Table 10. According to the results in Table 8, the SII values for the approach and landing tasks can be calculated using Equation (6). Given that the task sce-

TABLE 8: EAIs for PIFs.

PIF	EAI of the approach task	EAI of the landing task
Ground support	0	0
Crew workload management	-1.0	-1.0
Crew training and experience	-1.8	-1.8
Procedure format consistency and verification quality	-1.6	-1.6
Procedure quantity	0	0
Time stress	1.0	1.0
Adequacy of the man-machine interface	-0.6	-0.6
Adequacy of organization	0	0
Working conditions	1.6	1.6

nario does not change over the relatively short time required for the approach and landing phases, the SII value for both tasks is $\alpha = \sum_{i=1}^9 \theta_i = -2.4$. Finally, the HEP of each operation can be calculated based on Equation (7) as follows:

$$\text{HEP} = \text{CFP}_0 \times 10^{-0.504}. \quad (9)$$

The final results are presented in Table 9 and Table 10.

Based on hierarchical task analysis, the sequence structure of relations among approach operations is presented in Figure 3. Successfully setting the passenger sign does not affect the process of setting the landing light switches. Thus, the relation between O1.1 and O1.2 is considered to be independent. Then, the crew is required to set and crosscheck altimeters. O1.3 depends on the completion of O1.2. O1.4, O1.5, and O1.6 are parallel and independent. Consequently, considering the relations among these operations, the HEP of the approach task is calculated with Table 6. The result is obtained as follows.

$$\begin{aligned} \text{HEP}_{\text{approach}} &= \max \left(\left(\sum \text{HEP}_{\text{O1.1}} + \text{HEP}_{\text{O1.2}} \right), \text{HEP}_{\text{O1.3}}, \right. \\ &\quad \left. \left(\prod \text{HEP}_{\text{O1.4}} \times \text{HEP}_{\text{O1.5}} \times \text{HEP}_{\text{O1.6}} \right), \text{HEP}_{\text{O1.7}} \right) \\ &= 9.56E - 3. \end{aligned} \quad (10)$$

TABLE 9: HEPs for approach operations.

Operation	Cognitive activity	Cognitive function	Generic failure type	CFP ₀	HEP
1.1	Execute	Execution	E3	5.0E-4	1.57E-4
1.2	Execute	Execution	E5	3.0E-2	9.40E-3
1.3	Coordinate	Planning/execution	E5	3.0E-2	9.40E-3
1.4	Coordinate	Planning/execution	E1	3.0E-3	9.40E-4
1.5	Coordinate	Planning/execution	E1	3.0E-3	9.40E-4
1.6	Evaluate	Execution	E1	3.0E-3	9.40E-4
1.7	Coordinate	Planning/execution	E2	3.0E-3	9.40E-4

TABLE 10: HEPs for landing operations.

Operation	Cognitive activity	Cognitive function	Generic failure type	CFP ₀	HEP
2.1	Verify	Observation/interpretation	O1	1.0E-3	3.13E-4
2.2	Coordinate	Planning/execution	E5	3.0E-2	9.40E-3
2.3	Monitor	Observation/interpretation	O2	7.0E-2	2.19E-2
2.4	Verify	Observation/interpretation	O2	7.0E-2	2.19E-2
2.5	Verify	Observation/interpretation	O2	7.0E-2	2.19E-2
2.6	Evaluate	Execution	E1	3.0E-3	9.40E-4
2.7	Verify	Observation/interpretation	O3	7.0E-2	2.19E-2
2.8	Communicate	Execution	E4	3.0E-3	9.40E-4
2.9	Communicate	Execution	E4	3.0E-3	9.40E-4
2.10	Execute	Execution	E5	3.0E-2	9.40E-3
2.11	Verify	Observation/interpretation	O2	7.0E-2	2.19E-2
2.12	Execute	Execution	E5	3.0E-2	9.40E-3
2.13	Execute	Execution	E5	3.0E-2	9.40E-3
2.14	Execute	Execution	E5	3.0E-2	9.40E-3
2.15	Verify	Observation/interpretation	O2	7.0E-2	2.19E-2
2.16	Coordinate	Planning/execution	E5	3.0E-2	9.40E-3
2.17	Execute	Execution	E5	3.0E-2	9.40E-3
2.18	Coordinate	Planning/execution	E4	3.0E-3	9.40E-4
2.19	Verify	Observation/interpretation	I2	1.0E-2	3.13E-3
2.20	Monitor	Observation/interpretation	I2	1.0E-2	3.13E-3
2.21	Verify	Observation/interpretation	O2	7.0E-2	2.19E-2

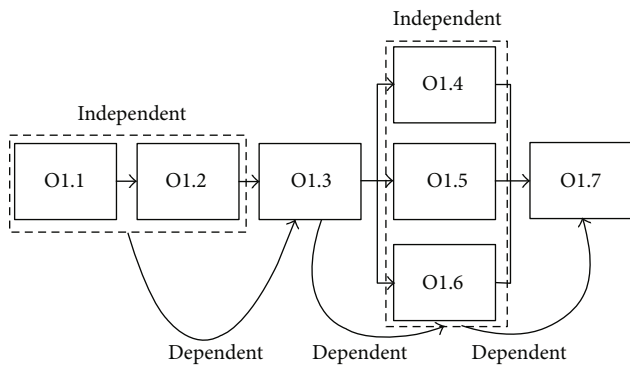


FIGURE 3: Relations among approach operations.

Similarly, the relation structure of the landing operations is presented in Figure 4. The HEP of the landing task is calculated as follows.

$$\begin{aligned}
 \text{HEP}_{\text{landing}} = \max & \left(\left(\sum \text{HEP}_{O2.1} + \text{HEP}_{O2.2} \right), \text{HEP}_{O2.3}, \right. \\
 & \cdot \left(\prod_{n=O2.4}^{O2.7} \text{HEP}_{O_n} \right), \text{HEP}_{O2.8}, K, \text{HEP}_{O2.17}, \\
 & \cdot \left. \left(\prod_{m=O2.18}^{O2.20} \text{HEP}_{O_m} \right), \text{HEP}_{O2.21} \right) = 2.19E-2.
 \end{aligned}$$

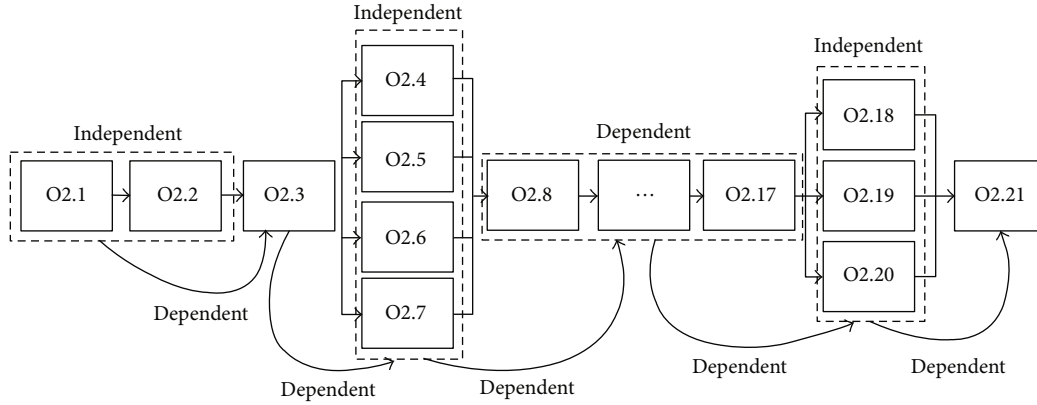


FIGURE 4: Relations among landing operations.

4. Results and Discussion

4.1. Comparison with the HEART Method. To demonstrate the validity of the proposed method, the quantitative results based on the modified CREAM method are compared with the results obtained with the HEART approach. Extensive experiments with the HEART method show that the approach HEP is $1.17E-2$ and the landing HEP is $2.51E-2$. The HEP values of approach and landing using different methods are listed in Table 11. The two calculations are not entirely based on the same data. Given the differences in the base probabilities, quantification process, and influence factors of these two methods, it seems inevitable that there will be differences in the results. The HEART model is a task-based method used to quantify HEP. This method mainly involves 8 generic tasks and 38 error-producing conditions. Analysts must select appropriate generic tasks and error-producing conditions for the specific operations in flight tasks to calculate the HEP. The HEART approach has been applied in various domains, such as nuclear plants, marine transportation, and the petrochemical industry [25–27]. Most studies consider the method reasonable and acceptable in terms of availability, validity, and consistency. However, in the quantification process of this method, the effects of external conditions on human cognitive functions, such as observation, interpretation, planning, and execution, are not adequately considered. These external conditions may also have overlapping effects on human performance. Furthermore, human behavior cannot be simply broken down into mechanical operations without considering cognitive processes because the operators must make more judgments during man-machine interactions than during routine operations.

In comparison with the HEART model, the proposed method is a cognition-based approach for quantifying human unreliability. Specifically, 9 PIFs are provided for flight scenarios to support human reliability analysis in flight tasks. In particular, the effects of these PIFs on human cognitive functions are fully considered in calculating the HEP for flight tasks. This approach may better represent the actual situation compared to the HEART approach, leading to more reasonable and safe results. In addition, the method can provide retrospective and prospective analyses to perform

TABLE 11: HEPs of the approach and landing operations.

HEP	Proposed model	HEART model
Approach	$9.57E-3$	$1.17E-2$
Landing	$2.19E-2$	$2.51E-2$

human reliability assessment, even if the results depend partly on expert judgment. Data scarcity is a major flaw in almost all current HRA methods, and the problem is difficult to solve. However, the proposed method is still an effective approach for assessing human error risk in complex flight tasks, as its inherent causal classification features can provide guidance for data collection and analysis.

4.2. Influence of Cognitive Functions on the HEP. In this section, the Pareto principle is introduced in human reliability assessment to identify the critical cognitive functions. The core concept of this principle is that 80% of the conclusions come from 20% of the causes [28]. Thus, this principle can assist engineers or analysts in optimizing cockpit ergonomic design and rationally allocating the cognitive load.

The hazard priority ranking of potential cognitive function failure is presented in Figure 5. The results indicate that the most important cognitive functions involve E5, O2, and E1, and the corresponding cumulative percentage is approximately 75%. Airlines should take some relevant measures for these three failure types to improve human performance and enhance flight safety. Action missed (E5) ranks first among the potential cognitive function failures. A long flight may cause a reduction in situation awareness for the crew, which can lead to misjudgments or even the omission of some operations. Thus, the flight crew should focus on adequately performing these operations. Adding intelligent auxiliary equipment may also be an effective measure. Notably, wrong identification (O2) is the second critical failure type that leads to human errors. Incomplete observations may be the main reason for target recognition errors due to the continuous interference of a complex external environment. Airlines should enhance crew training and decrease the interference effects for air traffic controllers and the crew in severe environments. Finally, the wrong type (E1) action is an important

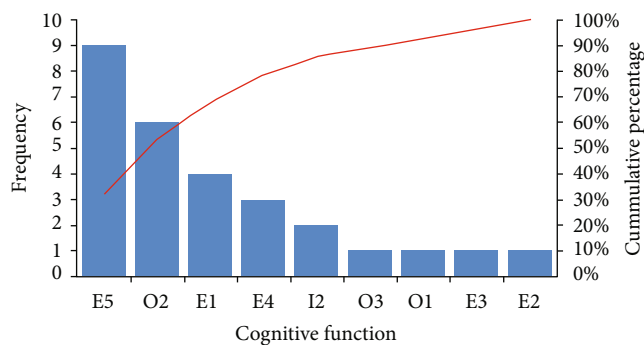


FIGURE 5: Pareto chart of cognitive function.

failure type that can increase the HEP. Facing a limited time during approach and landing phases, the crew is prone to make mistakes. Rational man-machine function allocation may reduce the operation time and workload of the crew and improve performance reliability.

5. Conclusions

In this paper, a modified CREAM method was developed and applied to human reliability assessment in flight tasks. First, a set of PIFs considering flight tasks, such as ground support and flight procedures, was constructed to reflect the causes of erroneous actions in flight. Then, the EAI and the SII were used to build a human reliability model. The qualitative descriptions of PIFs related to human reliability were quantified by the EAI. The SII reflects the influence of each PIF on the specific cognitive functions performed. Benefitting from the EAI and the SII, the computational cost is effectively decreased compared to that of the original CREAM method. Extensive experiments show that the modified CREAM method is a promising and effective approach for assessing human reliability in flight. Moreover, we provide some appropriate and reasonable measures for the most important cognitive function failure types to reduce human errors and improve flight safety. In future work, we will consider a dynamic human error assessment method with time sequences and physiological parameters.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (U1333119) and the National Defense Basic Scientific Research Program of China (JCKY2013605B002).

References

- [1] A. J. Erjavac, R. Iammartino, and J. M. Fossaceca, "Evaluation of preconditions affecting symptomatic human error in general aviation and air carrier aviation accidents," *Reliability Engineering & System Safety*, vol. 178, pp. 156–163, 2018.
- [2] D. D. Boyd, "A review of general aviation safety (1984–2017)," *Aerospace Medicine and Human Performance*, vol. 88, no. 7, pp. 657–664, 2017.
- [3] R. Westrum and A. J. Adamski, "Organizational factors associated with safety and mission success in aviation environments," in *Human Error in Aviation*, pp. 475–512, Routledge, 2017.
- [4] T. J. Etherington, L. J. Kramer, R. E. Bailey, K. D. Kennedy, and C. L. Stephens, "Quantifying pilot contribution to flight safety for normal and non-normal airline operations," in *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, pp. 1–14, Sacramento, CA, USA, 2016.
- [5] E. Hollnagel, *Cognitive Reliability and Error Analysis Method (CREAM)*, Elsevier, 1998.
- [6] Z. L. Yang, S. Bonsall, A. Wall, J. Wang, and M. Usman, "A modified CREAM to human reliability quantification in marine engineering," *Ocean Engineering*, vol. 58, pp. 293–303, 2013.
- [7] A. Wang, Y. Luo, G. Tu, and P. Liu, "Quantitative evaluation of human-reliability based on fuzzy-clonal selection," *IEEE Transactions on Reliability*, vol. 60, no. 3, pp. 517–527, 2011.
- [8] J. Calhoun, C. Savoie, M. Randolph-Gips, and I. Bozkurt, "Human reliability analysis in spaceflight applications," *Quality and Reliability Engineering International*, vol. 29, no. 6, pp. 869–882, 2013.
- [9] J. Calhoun, C. Savoie, M. Randolph-Gips, and I. Bozkurt, "Human reliability analysis in spaceflight applications, part 2: modified CREAM for spaceflight," *Quality and Reliability Engineering International*, vol. 30, no. 1, pp. 3–12, 2014.
- [10] J. Chen, D. Zhou, C. Lyu, and X. Zhu, "A method of human reliability analysis and quantification for space missions based on a Bayesian network and the cognitive reliability and error analysis method," *Quality and Reliability Engineering International*, vol. 34, no. 5, pp. 912–927, 2018.
- [11] B. Wu, X. Yan, Y. Wang, and C. G. Soares, "An evidential reasoning-based CREAM to human reliability analysis in maritime accident process," *Risk Analysis*, vol. 37, no. 10, pp. 1936–1957, 2017.
- [12] T. Bedford, C. Bayley, and M. Revie, "Screening, sensitivity, and uncertainty for the CREAM method of human reliability analysis," *Reliability Engineering & System Safety*, vol. 115, pp. 100–110, 2013.
- [13] S.-T. Ung, "A weighted CREAM model for maritime human reliability analysis," *Safety Science*, vol. 72, pp. 144–152, 2015.
- [14] A. C. Ribeiro, A. L. Sousa, J. P. Duarte, and P. F. Frutuoso e Melo, "Human reliability analysis of the Tokai-Mura accident through a THERP-CREAM and expert opinion auditing approach," *Safety Science*, vol. 87, pp. 269–279, 2016.
- [15] Q. Zhou, Y. D. Wong, H. Xu, V. V. Thai, H. S. Loh, and K. F. Yuen, "An enhanced CREAM with stakeholder-graded protocols for tanker shipping safety application," *Safety Science*, vol. 95, pp. 140–147, 2017.
- [16] E. Zio, "Challenges in the vulnerability and risk analysis of critical infrastructures," *Reliability Engineering & System Safety*, vol. 152, pp. 137–150, 2016.

- [17] M. C. DeHoff and S. K. Cusick, "Mental health in commercial aviation-depression & anxiety of pilots," *International Journal of Aviation, Aeronautics, and Aerospace*, vol. 5, p. 5, 2018.
- [18] T. B. Sheridan, "Human-robot interaction: status and challenges," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 58, no. 4, pp. 525-532, 2016.
- [19] C. Nan and G. Sansavini, "A quantitative method for assessing resilience of interdependent infrastructures," *Reliability Engineering & System Safety*, vol. 157, pp. 35-53, 2017.
- [20] X. Pan, Y. Lin, and C. He, "A review of cognitive models in human reliability analysis," *Quality and Reliability Engineering International*, vol. 33, no. 7, pp. 1299-1316, 2017.
- [21] X. He, Y. Wang, Z. Shen, and X. Huang, "A simplified CREAM prospective quantification process and its application," *Reliability Engineering & System Safety*, vol. 93, no. 2, pp. 298-306, 2008.
- [22] C. Airplanes, *Statistical Summary of Commercial Jet Airplane Accidents*, Boeing Commercial Airplanes, 2017.
- [23] K. A. Honn, B. C. Satterfield, P. McCauley, J. L. Caldwell, and H. P. A. van Dongen, "Fatiguing effect of multiple take-offs and landings in regional airline operations," *Accident Analysis & Prevention*, vol. 86, pp. 199-208, 2016.
- [24] S. Hirschberg, "Human reliability analysis in probabilistic safety assessment for nuclear power plants," *Safety and Reliability*, vol. 25, no. 2, pp. 13-20, 2005.
- [25] A. M. Kumar, S. Rajakarunakaran, and V. A. Prabhu, "Application of fuzzy HEART and expert elicitation for quantifying human error probabilities in LPG refuelling station," *Journal of Loss Prevention in the Process Industries*, vol. 48, pp. 186-198, 2017.
- [26] S. Rangra, M. Sallak, W. Schön, and F. Vanderhaegen, "On the study of human reliability in transportation systems of systems," in *2015 10th System of Systems Engineering Conference (SoSE)*, pp. 208-213, San Antonio, Texas, USA, 2015.
- [27] E. Akyuz and M. Celik, "A hybrid human error probability determination approach: the case of cargo loading operation in oil/chemical tanker ship," *Journal of Loss Prevention in the Process Industries*, vol. 43, pp. 424-431, 2016.
- [28] S. Liu, X. Gao, W. Fang et al., "Modeling the complex network of multidimensional information time series to characterize the volatility pattern evolution," *IEEE Access*, vol. 6, pp. 29088-29097, 2018.



Hindawi

Submit your manuscripts at
www.hindawi.com

