University of Strathclyde

Department of Naval Architecture, Ocean & Marine Engineering

Development of a human reliability assessment framework addressing specific needs and requirements of maritime operations

By

Sung Il Ahn

A thesis presented in fulfilment of the requirements for

the degree of

Doctor of Philosophy

Glasgow, UK

2022

Signed: Sung Il Ahn

Date: 4 August 2022

This thesis is dedicated to the memory of my father, Dae Joon Ahn, who loved me and presented me with lovely memories.

ACKNOWLEDGMENTS

Numerous individuals provided me with support and encouragement. I would like to express my gratitude to everyone who assisted me with my studies at this time.

First and foremost, I would like to express my sincere appreciation to my first supervisor, Dr Rafet Emek Kurt, who has been an excellent advisor during my research journey. I would also like to acknowledge the thoughtful and consistent assistance of my second supervisor, Professor Osman Turan.

Moreover, I would like to thank my colleagues for their support during my research, especially Dr Sefer Gunbeyaz, for his advice and help as a colleague and, sometimes, as a mentor.

I'd like to express my gratitude to my mother, Tae Yeon Kim, and my brother, Kyung Il Ahn, for their unwavering support, encouragement, and belief in me.

Last but not least, I'd like to thank my beloved wife, Ala Hwang, and my daughters, Sumin Ahn and Yeojin Ahn. Without the support and encouragement of my lovely wife and daughters, and their patience and understanding, I would not have reached the end of this journey.

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ABS	American Bureau of Shipping
AHP	Analytic Hierarchy Process
ATHEANA	A Technique for Human Error Analysis
BN	Bayesian Network
CAST	Causal Analysis based on Systems Theory
CFP	Cognitive Failure Probability
COCOM	Contextual Control Model
COG	Centre of Gravity
CPCs	Common Performance Conditions
CPT	Conditional Probability Table
CREAM	Cognitive Reliability and Error Analysis Method
CV	Crisp Value
DEMATEL	Decision-Making Trial and Evaluation Laboratory
EMSA	European Maritime Safety Agency
EPCs	Error Producing Conditions
ER	Evidential Reasoning
ERC	Emergency Release Coupling
ESD	Emergency Shutdown
FAHP	Fuzzy Analytical Hierarchy Process
FCMs	Fuzzy Cognitive Maps

FMEA	Failure Mode and Effects Analysis
FRAM	Functional Resonance Analysis Method
FTA	Fault Tree Analysis
FSA	Formal Safety Assessment
HEART	Human Error Assessment and Reduction Technique
HEP	Human Error Probability
HFACS	Human Factors Analysis and Classification System
HFE	Human Failure Events
HMI	Human Machine Interface
HRA	Human Reliability Assessment
HRAET	Human Reliability Analysis Event Tree
HTA	Hierarchical Task Analysis
IACS	International Association of Classification Societies
IMO	International Maritime Organization
ISM	International Safety Management
LNG	Liquified Natural Gas
MALFCMs	Marine Accident Learning with Fuzzy Cognitive Maps
MMI	Man Machine Interface
NASA	National Aeronautics and Space Administration
NARA	Nuclear Action Reliability Assessment
NOx	Nitrogen Oxides
PM	Particulate Matter
PSFs	Performance Shaping Factors

RBD	Reliability Block Diagram
SLI	Success Likelihood Index
SLIM	Success Likelihood Index Method
SMS	Ship safety Management System
SOLAS	International Convention for the Safety of Life at Sea
SOx	Sulphur Oxides
SPAR-H	Standardized Plant Analysis Risk Human Reliability Assessment
STAMP	System-Theoretic Accident Model and Processes
STPA	System Theoretic Process Analysis
THERP	Technique for Human Error Rate Prediction
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TTA	Tabular Task Analysis
UCA	Unsafe Control Actions
WMoM	Weighted Mean of Maxima

Abstract

The safety of life at sea is a top priority in maritime operations. Therefore, essential procedures and regulations are enforced to prevent loss of life. However, human error is one of the main contributors to accidents in safety-critical industries. For a human reliability assessment in the maritime domain, the main question is how we correctly understand the human factors in the maritime situation practically.

This research study aims to develop practical frameworks, which are different optimal combinations of analysis methods corresponding to the different research scope, to evaluate human errors in maritime operations that can more cleverly be identified, quantified, and integrated into the probabilistic risk assessment. Thus, achieving the aim is expected to improve overall safety within the maritime domain.

This research proposes four human reliability assessment frameworks corresponding to maritime systems' different complexities and interactions. Firstly, this study introduces the Bayesian CREAM framework, which is a method of determining the contextual mode for overall human error probability estimation. This method aims to determine the need for more specific HRA research and to support the quick analysis required by providing a simple and imminent calculation method. Second, the CREAM-based framework is developed to extend human errors to various human activities. This method minimises the expert's subjectivity while achieving a quantified human failure probability with a systematic and logical approach. Third, the SPAR-H-based framework is proposed to integrate human errors into a probabilistic risk assessment framework. This framework offers a new approach to human reliability, assessed through a customised reliability block diagram analysis to provide a new risk model. Finally, for a complex modern socio-technical system, a hybrid method is proposed to assess system reliability by combining the System Theoretic Process Analysis (STPA) and the Success Likelihood Index Method (SLIM). The proposed frameworks were applied to emergency response operations, including emergency steering, engine room fires, man overboard, and emergency shutdown system for the LNG bunkering process. The various frameworks established for human reliability evaluation will contribute to the wider utilisation of human reliability assessment, ranging from simple and convenient analysis performed by ship's crew

to complex analysis performed by expert human reliability analysts. Eventually, this study will enhance maritime safety by analysing human errors, identifying further problems, and adopting safety measures.

Contents

1 Introduction

1.1 Chapter overview

This chapter briefly introduces the background reasoning for initiating and fulfilling this research work.

1.2 General perspectives

Despite decades of technological advancements and the emergence of advanced systems-controlling capabilities, the accident rate has not decreased considerably, and new forms of accidents continue to occur. Human error is acknowledged as the primary cause of this problem. As a result, awareness of human error has evolved, and the solution appears to be changing individuals or positions within the system (Woods et al., 2017). On the other hand, people or work adjustments do not eliminate or diminish human error because human error is an unintentional phenomenon. This resulted in a new understanding that human error is more intimately tied to both systemic and personal problems (Dekker, 2017a). To reduce such human errors, it is required first to identify, quantify, and prioritise them. As a result, we started by calculating the probability of human errors to assess human reliability (Chauvin, 2011). Swain and Guttmann (1983) defined human reliability as the probability of humans doing the needed tasks within a specified time range. Human Reliability Assessment involves applying qualitative and quantitative approaches to ascertain the human component of risk (Bell and Holroyd, 2009).

Maritime accidents were viewed as a complex process (Abbaspour et al., 2020), with various components contributing to the accident's development. Human factors were also noted as significant in maritime accidents (Schröder-Hinrichs et al., 2013). Modern maritime systems are technologically advanced and incredibly reliable. However, because the maritime system is human-oriented, the rate of maritime fatalities continues to be high, manifesting itself in human fatalities (Rothblum, 2000). The maritime sector has had a significant number of disasters, and research indicates that most of these incidents are caused by human error (Chan et al., 2016). However, on the other hand, predicting human errors remains challenging due to the complexities and interactions of the human mechanism in the system and the unknown function of each component.

Furthermore, the advancement of new technologies in the maritime industry affects how the maritime industry responds to new challenges and opportunities. Numerous ships are being converted to automated control systems, either partially or entirely, employing cutting-edge technology. Human duties in cognitively intensive behaviours are becoming increasingly crucial in maritime, along with existing labour-intensive behaviours. As a result, human functions will become more difficult to define in safety management. Therefore, human reliability research must become a top priority for the maritime industry, and it cannot be delayed any longer and must begin soon. However, in light of the maritime industry's inherent safety-critical activities, this consideration of human reliability presents a significant challenge.

So, how can we depict a human interaction system to demonstrate the system's human error mechanism, assuming that we can accurately identify and quantify the human error? In that case, we can focus safety measures on preventing and mitigating the consequences of human error to reduce the number of incidents and thereby boost overall maritime safety.

1.3 Specific issues of human reliability assessment

It is challenging to comprehend the human mechanism by which humans respond to a specific scenario in safety-critical operations. As a result, this study introduces the HRA approach to represent this mechanism as understandable and interpretable. This section summarises and describes five distinct HRA-related issues discussed in greater depth in Chapter 2.

Firstly, Human Error Identification (HEI) is typically included in the Human Reliability Assessment (HRA), which assesses the system's influence on human error and error recovery (Kirwan, 1998). These HEI techniques range from the simple classification of errors to sophisticated software packages based on human performance models. However, anticipating human behaviour in complicated contexts is not an easy problem in and of itself, and human error identification 'technology' has much potential for improvement (Kirwan, 1992). By bringing an autonomous or software-controlled system into the maritime sector, human involvement in cognitively demanding behaviours and current labour-intensive behaviours will become increasingly necessary. While new technology can contribute to increased maritime safety, we should also consider human roles in complex systems such as autonomous software-intensive systems. This is because many human functions are difficult to define using traditional approaches such as a hierarchical task analysis, frequently utilised in present HRAs, in future complex systems. Although systematic techniques such as the System Theory Process Analysis (STPA) and the Functional Resonance Analysis (FRAM) have been developed better to understand the role of humans in complex systems, the HRA applications have not yet been systematised. As a result, assessing what enhancements have been made to the existing HRA to adapt it to the changed or added human function is required.

Secondly, the nominal human error probability data would be the foundation of human reliability theory and practice. However, throughout the history of human reliability assessment, it has been exceedingly difficult to gather and provide meaningful and available data (Kirwan et al., 1990). Thus, analysts are limited in conducting human reliability assessment (HRA) due to a lack of raw figures on human errors. This condition may amplify risk analysis volatility and degrade the effectiveness of HRA outcomes (Liu and Li, 2014). As a result, when anticipating human error in a particular context, most HRA approaches, such as THERP, HEART, and SLIM, rely on expert judgement perspectives (Svenson, 1989). This may introduce ambiguity into the results when evaluating Performance Shaping Factors (PSFs) affecting human performance and failure. Additionally, efforts should be made to eliminate bias and gain consensus among experts.

Thirdly, it has been stated that the HRA invested great effort in building sophisticated representations of human performance through an ever-expanding list of performance shaping factors (Boring, 2010) to develop a secure method that considers all possible performance shaping factors (Boring et al., 2007). However, determining the optimal number of PSFs employed, dependency, the relative significance of each PSF, and the ranking of individual PSFs are problems that must be resolved. Additionally, when applying this PSF perspective to the maritime industry, another issue is that the developed PSFs are based on the nature and environment of the nuclear and aviation sectors, not on human activities in the maritime sector.

The next problem is to incorporate human error into probabilistic risk assessment. The limitations of human error quantification have long been recognised as constraints when using probabilistic risk assessment (Cooper et al., 1996). The probabilistic risk of human error should not be determined solely by the failure probability of a single human error. Instead, it should consider recovery action and dependency, among other components. However, current task modelling representations lack the expressive strength necessary for a systematic and exact description of possible human errors (Fahssi et al., 2015).

Finally, there are no obvious criteria for determining the appropriate HRA method to use given the study scope, complexity, and interaction of the system. As a result, there is a risk that excessive effort will be spent or insufficient findings will be acquired throughout the study. Furthermore, even when appropriate HRA methods are chosen, not all HRA methods provide a comprehensive framework for identifying, quantifying, and modelling human error, necessitating the development of a diverse range of frameworks depending on the degree of complexity and interaction of a given system.

1.4 The layout of the research study

This chapter has presented some background to the issue addressed in this thesis. Therefore, this research study will investigate the contribution of HFs to past maritime accidents, aiming to develop a resilience engineering approach to improve maritime safety. ***Figure 1. 1*** provides the overall layout of this research study. Thus, the structure of this thesis can be summarised as follows:

- Chapter 1 provides background information and discusses the requirements for designing this research. This chapter aims to clarify and emphasise the motivation for this research by addressing existing issues and problems in human reliability assessment.
- Chapter 2 critically evaluates the prior research on human resource assessments. The existing literature in various industries is investigated to complete this thesis. Additionally, this chapter discusses the theoretical foundations of several human reliability assessments and includes practical examples. Finally, this chapter aims to highlight research gaps in the application of human reliability management to the maritime industry.

- Chapter 3 discusses the rationale for this research. Additionally, it describes the research questions that will be addressed, including the aim and objectives of this research.
- Chapter 4 describes the research approach and methodology used in this study.
- Chapter 5 details the process of selecting the most appropriate HRA technique for the case study. Fourteen criteria are established to compare HRA methods, and four independent evaluations are undertaken based on the TOPSIS method. Finally, the optimal combinations of HRA's techniques for assessing human reliability are selected for each case study described in Chapters 6–9.
- Chapter 6 explains how to determine the human error mode by analysing the context in which human actions are performed. The Bayesian CREAM approach estimates the overall human error probability through human error mode determination. Finally, the chapter illustrates a case study for emergency steering operations.
- Chapter 7 discusses a technique for analysing extended human errors. Individual human error probabilities are evaluated for human tasks required to complete emergency operations, including cognitive activity, using a hybrid Fuzzy and CREAM methods. A case study of a ship's engine room fire drill was demonstrated to illustrate the practical application.
- Chapter 8 provides a SPAR-H-based framework to determine the probability of human error in a procedure where humans and machines interact. The system reliability diagram is used to recreate the derived human error probability for incorporation in the probabilistic risk assessment. The developed framework for analysing human reliability is used to conduct case studies on man overboard emergency training.
- Chapter 9 describes the combined method of the STPA and the SLIM, which is a comprehensive framework. The framework employs the STPA method to identify human errors within complex systems, and the SLIM is used to quantify the identified human errors. The reliability block diagram is used to model the probability of human errors to undertake system reliability evaluation. LNG bunkering's emergency shut-down system is selected in the case study.
- Chapter 10 summarises the study's major findings, discusses the study's limitations, and makes recommendations for further research.

- Chapter 11 presents the conclusions of this research study.

The layout of this thesis is depicted in ***Figure 1. 1***.

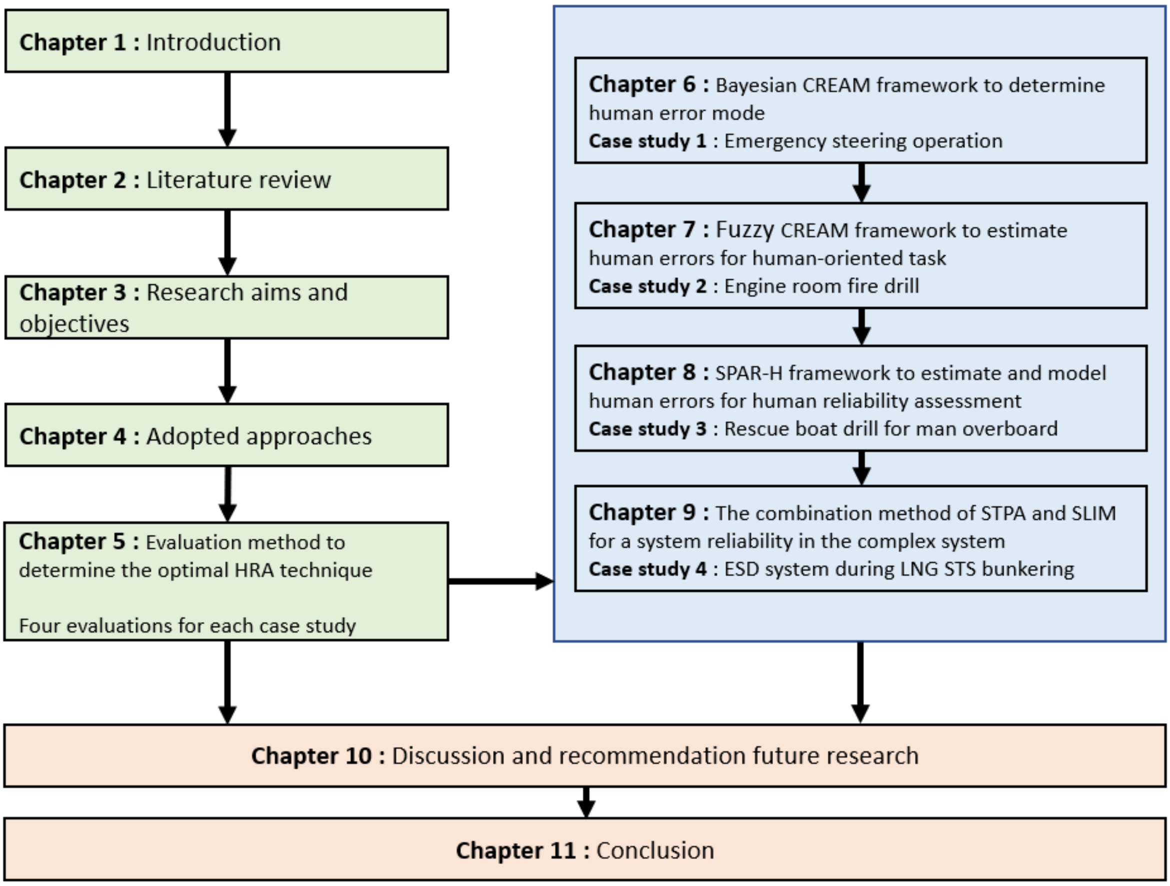

Figure 1. 1 The layout of the research study

1.5 Main contributions and novelties of the research study

Previous research has demonstrated that the human factors of ships operating in the maritime sector significantly contribute to accidents. However, it is not easy to ascertain the mechanisms by which human activities result in human error. The complexity of operations and systems on ships and offshore platforms varies, as does the degree of interaction between humans and humans, as well as humans and technology (hardware-software). The majority of HRA techniques are geared toward quantifying human error. Each HRA method is created for a specific purpose, and the fact that many existing HRA methods do not completely cover the qualitative and quantitative stages should not be interpreted as a deficiency of those methods. However, it is critical to developing an optimum framework by combining appropriate techniques for each analysis phase to evaluate human and system reliability for various tasks. The human reliability assessment must also be completed using the HRA approach most relevant to the research scope. In this context, this study proposes four different frameworks to address various analysis capacities depending on the complexities and interactions of operations. Then, processes and systems in real-world maritime are analysed and depicted to demonstrate the suggested framework's feasibility. The criteria and evaluation process for determining the most suitable approach to human reliability is also provided. The established frameworks will likely boost human reliability assessment in the maritime domain, contributing to increased safety. The novelty of each framework is discussed in more detail below.

- This research introduces a method of determining the contextual mode for overall human error probability estimation using the BN-CREAM method.
  - ➢ The BN-CREAM method aims to determine the need for further specific HRA research and to support the urgent analysis required by providing a simple and imminent calculation method. In addition, this method enables reflection of the specific features of the Maritime Operation to human performance through assessment of the scenario and procedures of emergency steering operation to evaluate the impact of particular characteristics of maritime operations on human performance and identify significant factors among them.

- This research study provides two CREAM-based frameworks to assess human errors through experts' judgement. This method reduces the expert's subjectivity while achieving a quantified human failure probability with a systematic and logical approach.
  - The novelty of this model is that this method utilises the Fuzzy membership to obtain quantified results but does not use the Fuzzy logic avoiding the loss of useful information by the 'If-Then Rule' of this Fuzzy logic. This is an innovative attempt to utilise the nature of the Fuzzy membership during the Common Performance Conditions (CPCs) weighting process and consider interdependence. In addition, this study provides a novel framework consisting of scenario assessment and onboard procedures verification, which intends to provide a framework for recognising the form in which we may evaluate practical human behaviour and quantify the human error. For illustration purposes, a case study on emergency response during an engine room fire drill was conducted.
- This research study provides the SPAR-H based framework to assess human reliability with a scenario-oriented approach for the maritime emergency drill.
  - Notwithstanding that, only a few human reliability studies are particularly applied to shipboard emergency drills in the maritime industry. The new aspect of the proposed hybrid method is that it combines SPAR-H and Fuzzy sets with a custom task analysis technique. In addition, this study developed a modified system reliability block diagram that enables the calculation of dependencies between system components and obtains entire system reliability from the error probability of each element. A case study on the men overboard onboard ships was carried out.
- A novel hybrid method for assessing system reliability is developed by combining the System Theoretic Process Analysis (STPA) and the Success Likelihood Index Method (SLIM).
  - This study developed a new hybrid method combining the STPA and the SLIM to analyse human duties qualitatively and quantitatively for safety-critical complex operations. This systemic approach based on the STPA was created to assist in understanding human process models and capturing additional causal scenarios. The human process model with PSFs is unique as it proposes a new simplified

model of the human diagnosis process from a system perspective. The development process of accident scenarios is a newly proposed guideline that can be quickly applied to identify a rich set of scenarios related to human behaviour, including system information, human diagnosis processes, and performance shaping factors. The advantage of human task analysis by the STPA is that it provides a muti-dimensional analysis for each context of human responsibility, not a collective analysis according to the form of human action by general decomposition. Furthermore, a new human controller process model is developed, which provides more detailed information for accident scenarios and expands it to elements outside the system that affect human performance. The proposed framework provides a structure for both qualitative and quantitative risk assessment.

1.6 Research outputs

The following publications resulted from this research study.

1.6.1 Journal papers

- **Sung Il Ahn** and Rafet Emek Kurt (2020). Application of a CREAM based framework to assess human reliability in emergency response to engine room fires on ships. Ocean Engineering, pp1-15. https://doi.org/10.1016/j.oceaneng.2020.108078
- **Sung Il Ahn**, Rafet Emek Kurt and Emre Akyuz (2022). Application of a SPAR-H based framework to assess human reliability during emergency response drill for man overboard on ships. Ocean Engineering, pp1-14. https://doi.org/10.1016/j.oceaneng.2022.111089
- **Sung Il Ahn**, Rafet Emek Kurt and Osman Turan (2022). The hybrid method combined STPA and SLIM to assess the reliability of the human interaction system to the Emergency shutdown system of LNG ship-to-ship bunkering. Ocean Engineering. (Under review)

1.6.2 Conference papers

- B. Navas de Maya, **Sung Il Ahn** and Rafet Emek Kurt (2019). Statistical analysis of MAIB database for the period 1990-2016. International Maritime Association of the Mediterranean (IMAM), Annual Congress, 2019, Varna. https://doi.org/10.1201/9780367810085-67

1.7 Chapter summary

This chapter summarised the general reasons for pursuing this research study, including identifying a gap regarding how to approach the human element systemically in maritime accidents. It also summarised the layout of this thesis and provided a diagram in ***Figure 1. 1***, which allows a smoothly reading flow. Chapter 2 provides a critical literature review. Moreover, Chapter 2 also identifies the research gaps that will be addressed within this research study.

2 Literature review

2.1 Chapter overview

A critical review of the existing literature is carried out in this chapter, which is geared toward covering the numerous areas of interest that the researcher selected to finish this thesis.

2.2 Human error and human factors

Safety is a critical concern in the maritime industry. Still, it remains a difficult task to forecast and avoid accident occurrences since the reasons for failure are complex and include a variety of elements. Particularly concerning is that human factors in ship operation significantly contributed to the accident. Human error, for instance, has been identified as a primary contributing factor to maritime accidents, with estimates ranging from 65 to 90 per cent. (Kristiansen (2013); Ung (2015); Akyuz et al. (2018); Kurt et al. (2016b); Antão and Soares (2019)). On the other hand, human factors and errors are frequently used without explicit knowledge of what they mean (Khan, 2008). This section provides a quick overview of the significance of human error and the human factor.

The human error refers to something that was "not intended by the actor; not desired by a set of rules or an external observer; or that led the task or system outside its acceptable limits" (Senders and Moray, 1991). Several classifications are used to categorise human errors in HRA research. According to Rigby and Franks (1970), human error is defined as any member of a collection of human acts that exceeds a certain threshold of acceptability. As a result, an error is just an activity that occurs outside the system's tolerance range, where the system sets the bounds of tolerable performance. According to Swain and Guttmann (1983), a human error can be defined as "any individual action or member of a group of activities that exceeds some limit of acceptability". The Health and Safety Executive (Books, 2009) defines human factors as environmental, organisational and job elements as well as human and individual characteristics that influence behaviour at work in a way that can harm health and safety at work. The Swiss Cheese Model, first defined by Reason (1990), methodically explained the human aspects.

As indicated in ***Figure 2. 1***, human factors are categorised into organisational effects, supervisory factors, preconditions and unsafe acts. However, no causal categories for each measure were supplied to distinguish between active and latent failure. To do this, the HFACS created a more systematic way of classifying human factors, as seen in ***Figure 2. 2***. Performance shaping factors (PSFs) are used to adjust the nominal human error probability based on the detailed categories of human factors. Furthermore, because conditions influencing human performance have been analysed in terms of several context factors, the detailed types of human factors are also useful in predicting human errors (Lee et al., 2011b).

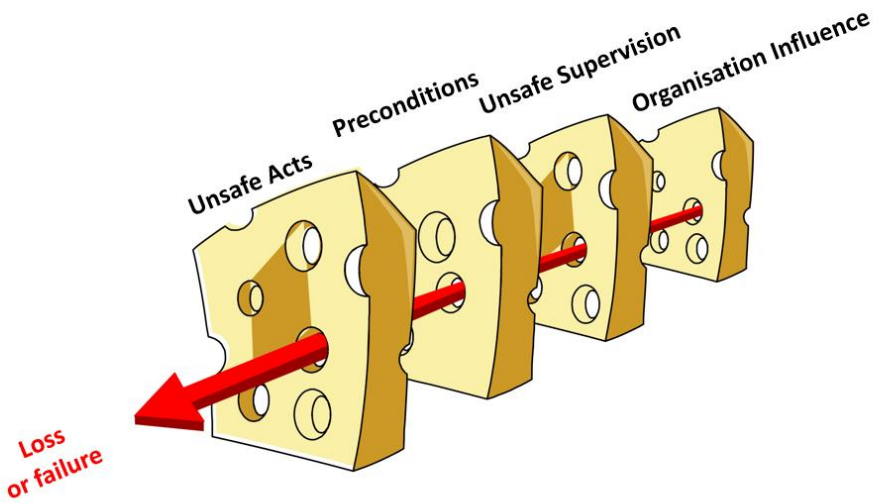

Figure 2. 1 Swiss cheese model (Adopted from Park (2018))

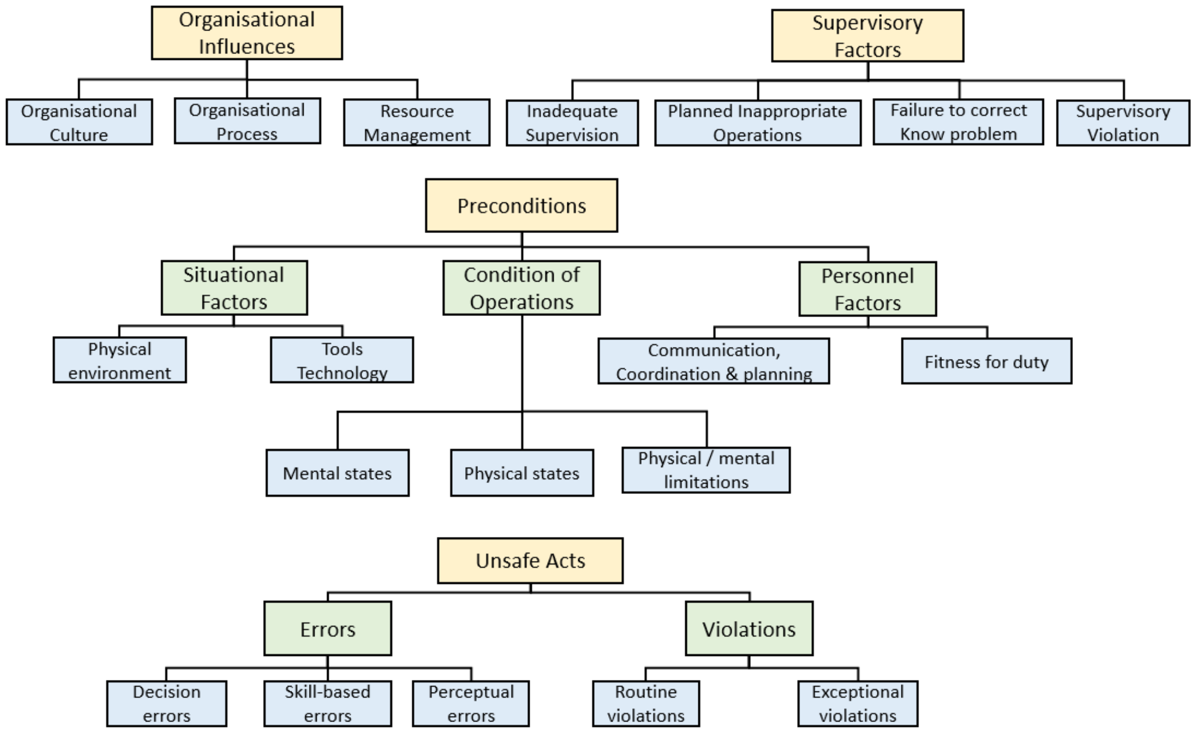

Figure 2. 2 Human Factors Classification (Adopted from Shappell and Wiegmann (2000))

2.3 Human reliability assessment

Human Reliability Assessment (HRA) is a comprehensive and systematic technique that employs qualitative or quantitative methodologies to evaluate human contributions to the reliability of a system. HRA also employs a variety of methods to assess how human performance can negatively impact a system, estimate the frequency of its occurrence, and determine its significance. HRA is based on the implicit assumption that "human error" is a meaningful concept that can be linked to individual behaviour (Hollnagel, 2005). HRA approaches are typically classified into two categories: those that use a database and those that use expert opinion. The approaches in the first category have a nominal human error probability database. Once these nominal error probabilities have been manipulated, the assessor can extrapolate the generic information to the individual case under consideration. This manipulation is often based on the assessor's context-related judgement (Kirwan, 1996). Due to the lack of human error data, most available HRA methodologies rely on expert judgement techniques (Musharraf et al., 2013). Additionally, human reliability is critical to the total system reliability perspective. This shows that individuals will likely execute as planned within a specific period of time and under specified environmental circumstances (Sgobba, 2017). Various ways have been developed over the last half-century to achieve the present third generation HRA methods via first generation HRA methods (De Felice and Petrillo, 2018). The first HRA approach combined quantitative risk analysis with human behaviours and errors. The PSF is found once the tasks are broken down into discrete activities. However, when looking at the PSF, these methodologies overlook the cognitive element, errors of commission, context, and organisational issues. The Technique for Human Error Rate Prediction (THERP) and the Standardized Plant Analysis Risk Human Reliability Assessment (SPAR-H) are two first-generation HRA methodologies examined in this study. In the second generation of human error techniques, the error of commission, context, and cognitive processes are all considered. It is more complicated since it focuses on the cognitive aspect of human reliability. In contrast, the previous generation HRA just looked at the behavioural aspect. There is more psychology involved in the second generation HRA, such as the CREAM and the ATHEANA. The HRA of the third generation is a hybrid method of the first and second generations. As a result, first generation approaches have been rebranded, for example, the revised Human Error

Assessment and Reduction Technique (HEART) has been called Nuclear Action Reliability Assessment as a third generation technique (NARA).

2.4 The overall HRA process

The overall process of HRA will be described in this section, as shown in ***Figure 2. 3***, and the essential components will be briefly defined. Because various HRA techniques have distinct ways of completing the qualitative and quantitative parts of the research, section 2.5 explores more into the specific HRA methods procedure.

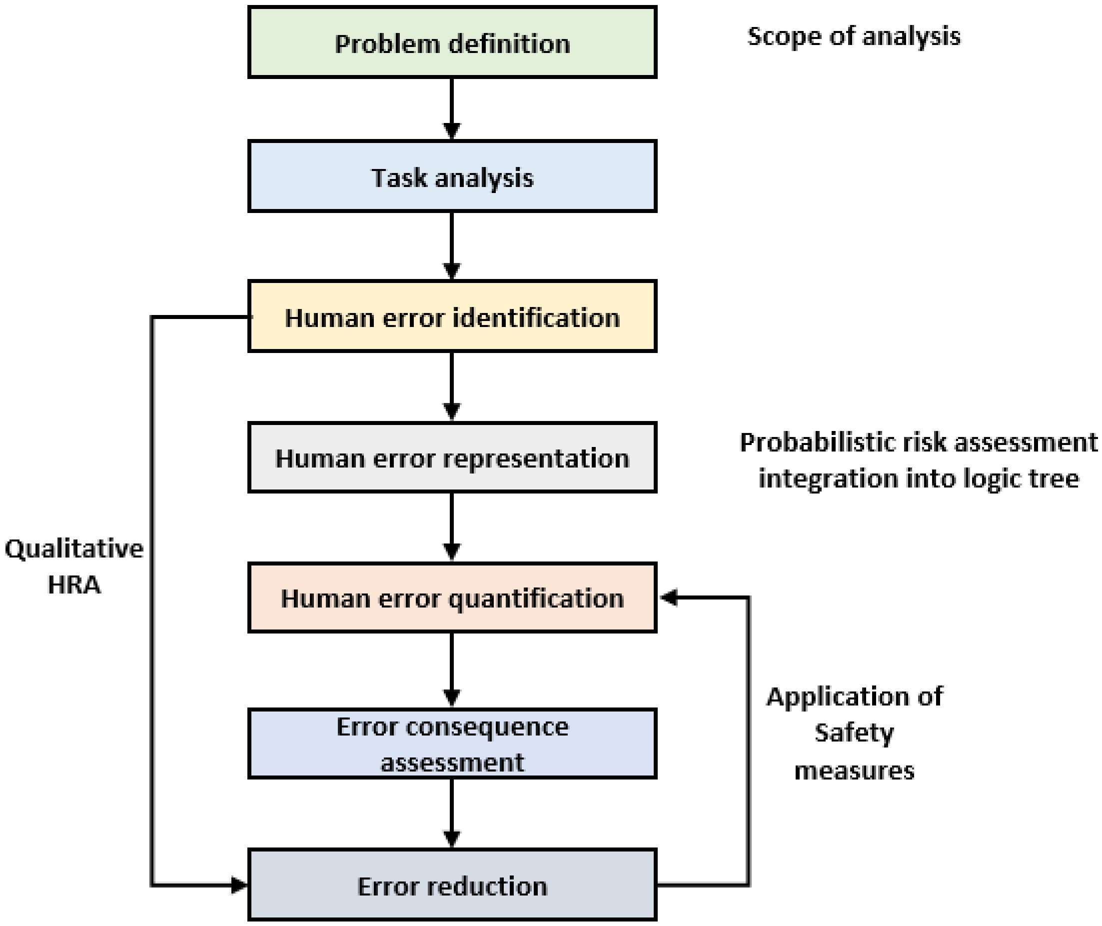

Figure 2. 3 Overall HRA process

2.4.1 Problem definition

The term "problem definition" refers to the process of identifying the human tasks or responsibilities that will be evaluated (Kirwan, 1994). This might be stated as a system's definition or as the operational procedure to be assessed. This step is to establish the study's scope. This procedure determines whether to address only operational failures or also maintenance, monitoring, and other activity failures. Additionally, it should be established at this point if the analysis will be quantitative, qualitative, or a combination of both. Thus, the problem definition is the initial phase in risk assessment in which policy objectives, scope, evaluation endpoints, and procedures are condensed into an officially defined strategy for problem analysis (Wolt et al., 2010).

2.4.2 Task analysis

Task analysis represents the understanding of factors influencing human performance and the information requirements of system designers (Annett and Stanton, 2000). Once the human aspect of the problem has been described, task analysis can be used to determine what human actions should occur in such situations and what equipment and other interfaces the operators should employ. However, a system is a complex collection of interconnected components that might contain humans and machines. In addition, these components interact to provide a purpose (Shepherd, 2003). Consequently, task analysis is required to examine how tasks are performed, including detailed descriptions of manual and mental activities, task and element durations, task frequency, task allocation, task complexity, environmental conditions, required clothing and equipment, and other unique elements associated with or required by one or more individuals (Ainsworth and Kirwan, 1992). In addition, this process may reveal key human aspects in tasks and the possibility of human error (Swain and Guttmann, 1983). Task analysis, also known as job decomposition, is a systematic method that identifies and decomposes each task into the steps and sub-steps that make up the human activities necessary to achieve a system's objective. (Annett, 2003).

The most commonly used task analysis approach is hierarchical task analysis (HTA), which is particularly beneficial for assessing human reliability since it is simple to implement yet highly effective at resolving a wide variety of difficulties (Kirwan, 1994). The HTA requires the analyst to describe a task in terms of a hierarchy of operations and plans (Shepherd, 1985). The HTA has endured representing a system sub-goal hierarchy for extended analysis. The HTA has been utilised for various purposes, including interface design and evaluation, function assignment, task support design, error prediction, and workload evaluation (Stanton, 2006). One element of HTA that creates a simple method for task analysis is that it uses a single primary analysis representation(Diaper and Stanton, 2003).

2.4.3 Human error identification

The human error refers to behaviours that can result in a degraded system state, either alone or in conjunction with other hardware or software failures or external events and should be considered during risk analysis. By examining the types of system failures humans contribute to, design principles that limit the occurrence and effect of errors can be established (Norman, 1983). Following the completion of the task analysis, the identification of human error is the next phase. Human Error Identification (HEI) is a specialised area of human factors to predict the types of errors that may arise by analysing tasks and the characteristics of the technology (Stanton and Baber, 1996). The development of appropriate preventative and/or mitigation methods may be made possible by identifying the errors that frequently lead to incidents and accidents (Baysari et al., 2009). Thus, the analyst evaluates what may go wrong and the root cause of the failure at this step (Kirwan, 1994). Identifying errors begins with establishing the scope of the analysis (Kirwan, 1998). The analysis's objective in this phase is to identify which operator will intervene, for example, whether to include just emergency events, misdiagnosis, maintenance faults, or rule violation errors. Depending on the analysis technique employed, human errors can be categorised in a variety of ways. Commission errors, missing errors, timing problems, and period errors are examples of human error modes. In addition, the causal factors of these human errors vary, such as various controller errors related to recognition, memory, decision-making, communication, and team resource management (Shorrock and Kirwan, 2002). The conventional strategy for identifying human errors is decomposing systems or goals until human behaviour is understood. The primary purpose of applied human reliability analysis is to quantify human error, with minimal focus placed on preventing human error. Although decomposition is an effective and easy tool for identifying human errors, it has some disadvantages. To begin with, developing safety measures for human errors is difficult since the literature does not give adequate data on the causes of human errors. Second, it is difficult to accurately depict two distinct facets of humans: those who produce errors and those who experience recovery failures. Finally, no relationship between human error and other software, hardware, or environmental events is modelled. Chapter 9, research on system reliability, including human error, in a Complex System, will cover this subject in depth.

2.4.4 Human error modelling/representation

Human error representations are often referred to as "modelling" because they facilitate the nonverbal delivery of data, relationships, and conclusions. Human error modelling aims to calculate a system's total risk level by combining the probability of all failures and the combination of failures (hardware, software, humans, and the environment). For example, the ability to predict the likelihood of human error allows system designers to modify system designs to achieve higher quality results and information about required skill levels, training programs, task design, task assignment, and work organisation (Elmaraghy et al., 2008). In addition, Fahssi et al. (2015) argue that systematic capturing of human errors within a working model helps design and evaluate error-tolerant interactive systems. The modelling of human errors is achieved by representing human errors with defects or other errors in the event tree naming logic trees (Kirwan, 1994). A typical example is Fault Tree Analysis (FTA), a top-down logical approach to fault analysis. Unwanted situations in the system are analysed using Boolean logic to combine a series of sub-level occurrences (Lacey, 2011). In human reliability analysis, the fault tree describes the set of human errors and their impact on system goals.

Meanwhile, event tree analysis (ETA) is an established risk analysis technique for assessing the likelihood of accidents in a probabilistic context(Ferdous et al., 2009). Event trees are usually binary logical trees that evolve from initial events to a possible result or logical set of results from the system, but exceptions exist. Nevertheless, the main goal of human error modelling is to discover how the probability of failure or success in one task relates to the likelihood of failure or success in another (Swain and Guttmann, 1983). Various industries have researched such human error representation. For example, Gosling et al. (1998) studied the aviation industry to improve the description of human errors to identify the form of aviation accidents and develop preventive methods for human errors that cause accidents. On the other hand, in the maritime industry, Akyuz and Celik (2014) applied modelling of human errors to various maritime disaster cases to analyse human factors' involvement in the event process. In Chapter 8, this study suggests and explores a unique approach based on a system reliability block diagram for human error representation.

2.4.5 Human error quantification

Human error quantification is a phase in which the probability of human error is quantified, and subsequently, the overall influence on system safety or dependability is determined. The human error probability (HEP) is defined as follows.

$$\mathrm{HEP} = \frac{\text{The number of times an error has occurred}}{\text{The number of opportunities for error to occur}} \qquad (2\text{-}1)$$

There may be two ways to determine the probability of human error: an approach based on data and an approach based on expert judgement. Alternately, these methods can be separated into task-based and context-based categories (Zhiqiang et al., 2009). However, due to the scarcity of human error data, most HRA research primarily relies on expert opinion (Dekker, 2017b). Expert judgement estimates human errors by converting the evaluation of variables influencing human performance, known as Performance Shaping Factors (PSFs), into numerical multipliers (Boring et al., 2007). As seen in ***Figure 2. 4***, once the nominal human error value is calculated, it is calibrated to account for the positive or negative impacts of PSFs. Since the method for estimating human errors varies according to the HRA used, specifics will be presented in Section 2.5.

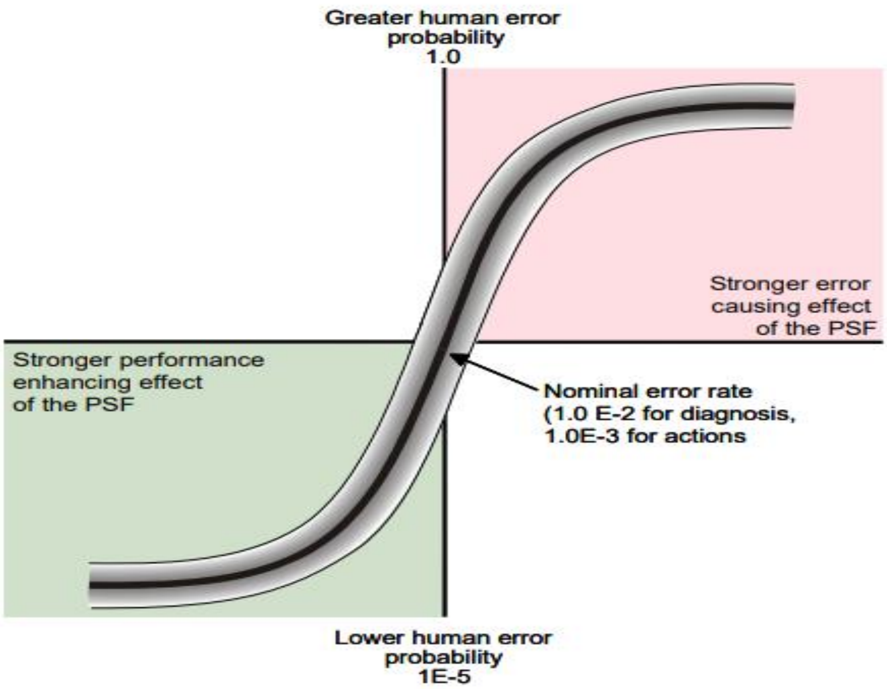

Figure 2. 4 Function of Human error probability based on PSF (Gertman et al., 2005)

2.4.6 Impact assessment

When the probability of human error (HEP) is assigned to various events of a failure or event tree (or both), these trees are quantitatively evaluated (Kirwan, 1994). Here, it is possible to investigate and analyse the frequency and uncertainty distribution of undesirable outcomes and/or risks (Cacciabue, 2004). This step allows for the determination of the system's overall risk level. Then it may be evaluated whether the system presents an acceptable degree of risk. For example, suppose human error substantially contributes to the system's risk level, and the amount of systematic risk is determined to be excessive. In that case, the appropriate errors will be targeted for error reduction.

2.4.7 Error reduction analysis

The human error reduction phase is a procedure for boosting the overall system's safety and reliability by strengthening the safety layers for discovering human errors. Redundancy is a general approach for enhancing system reliability (Senders and Moray, 2020). The human error reduction strategy is judged differently depending on whether HRA is conducted quantitatively or qualitatively. For instance, the qualitative analysis might fundamentally obstruct the indicated error route, logically preventing errors. By contrast, quantitative analysis may be used to adopt safety measures and then re-quantify them to evaluate if they should be tolerated. However, determining the amount to which the error reduction technique should be accepted while undertaking qualitative analysis is problematic. For example, while taking safety precautions for all recognised faults would be desirable, this may not always be achievable. The assessors' particular needs determine the error reduction tactics. Even if the qualitative review of human error is necessary, quantitative assessment should be considered to identify the nature of the error, at least in part, when an overall risk assessment of the system is required (Embrey, 2004).

2.5 Review of certain HRA techniques

The preceding section briefly outlined the general process of HRA. This section goes into further depth about each HRA approach.

2.5.1 THERP

The Technique for Human Error Rate Prediction (THERP) technique, developed in the Sandia Laboratories for the US Nuclear Regulatory Commission, is a well-known tool based on the event-tree approach for evaluating the probability of a human error in a manner comparable to an engineering risk assessment (Swain, 1964, Swain and Guttmann, 1983). This technique also takes into consideration performance shaping factors (PSFs), which are factors that may have an impact on this probability. This approach uses a database of error probabilities updated by PSF and other parameters (Kirwan, 1996). The database created by Swain and Guttmann using simulators and accident reports is available online. THERP is an iterative technique that consists of the four phases outlined in the following section. Although not necessarily in the same order, this is done repetitively until the system deterioration caused by human error has been reduced to all tolerable levels (Swain, 1964). As a technical component, the human is deemed as such, and the generated tree depicts the phases involved in a task in the sequence in which they should be performed (Castiglia et al., 2015). THERP is a comprehensive methodology that deals with task analysis, error identification and representation, as well as the estimation of human error probability (HEP). This subsection primarily discusses a strategy for quantifying human error and briefly explores the dependency model that is used in conjunction with a human-reliability-analysis event tree (HRAET). The outline of a THERP procedure is illustrated in ***Figure 2. 5***.

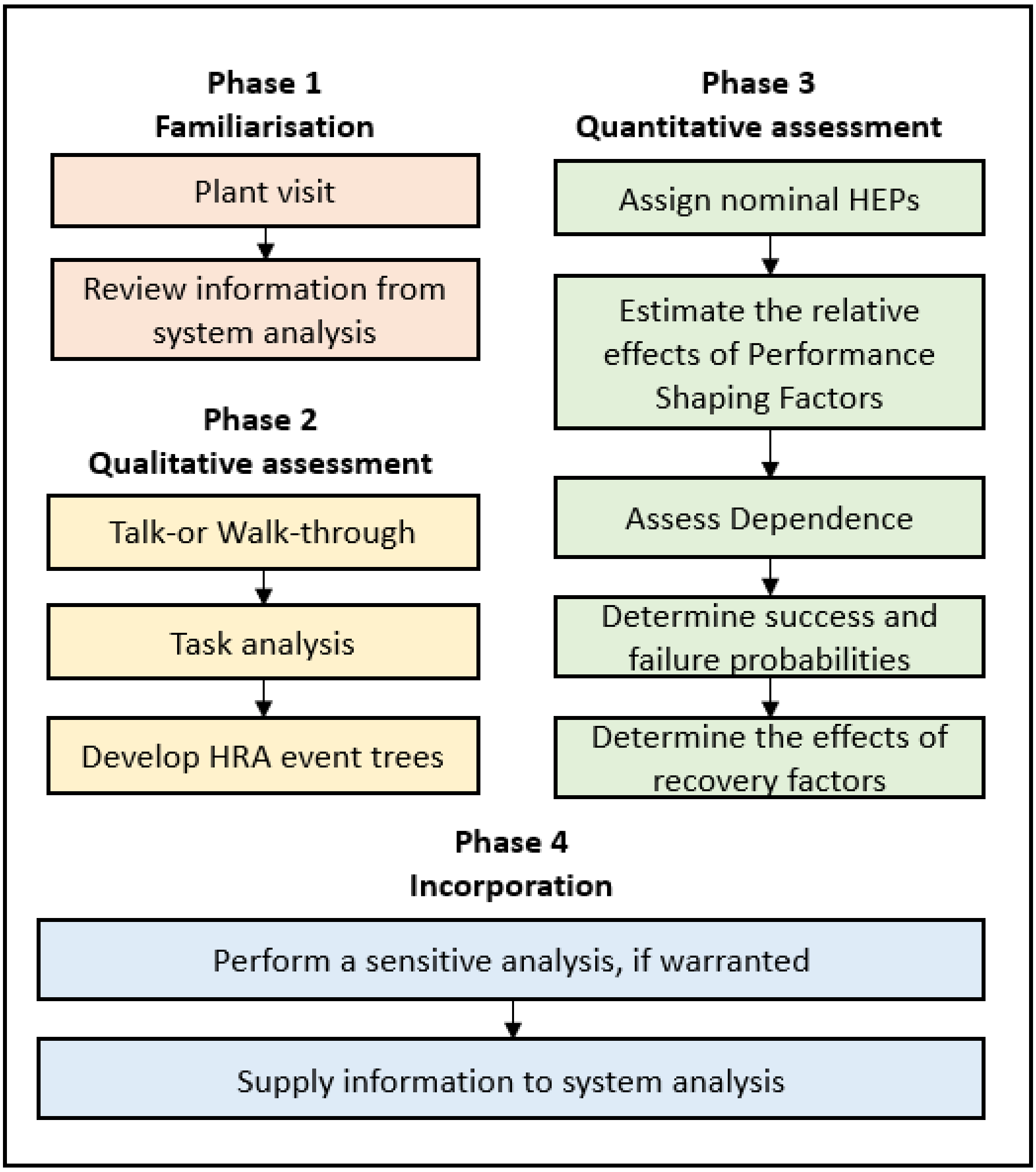

Figure 2. 5 Outline of the THERP procedure for the HRA(Swain, 1983)

The following are the three distinct aspects employed in the Human Reliability assessment in THERP.

1) A human-reliability-analysis event tree (HRAET)

A graphic task analysis method called the HRA event tree is presented to diagramming correct and incorrect human actions, as shown in ***Figure 2. 6***.

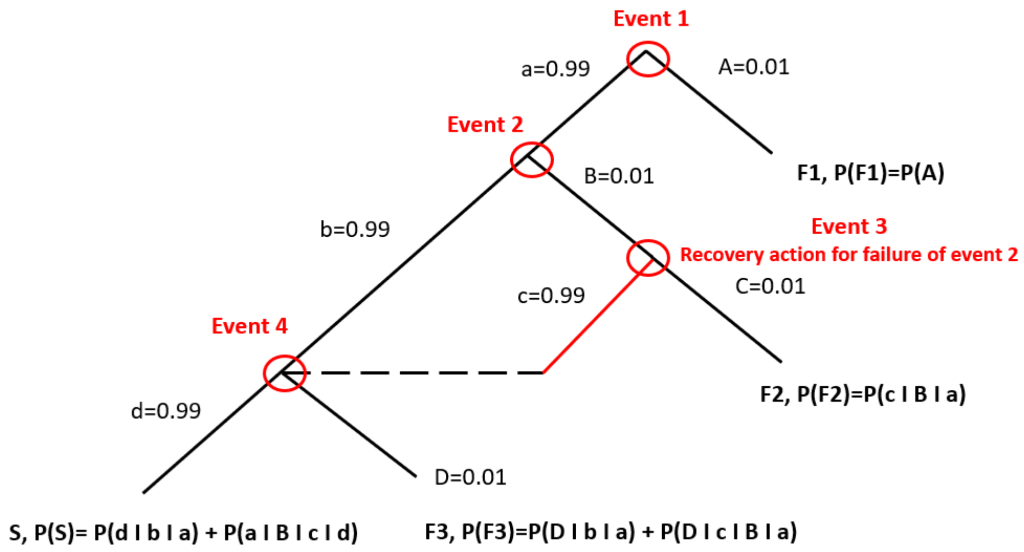

Figure 2. 6 Human Reliability Analysis Event Tree(Adopted from Swain (1964))

2) Dependency

There is a spectrum of possible levels of dependency between complete dependence and zero dependence. A dependence model of THERP in which the infinite possible levels of dependence are represented by the following five levels: zero dependence (ZD), low dependence (LD), moderate dependence (MD), high dependence (HD), and complete dependence (CD). The conditional probability for" Task N", given the success or failure probability of the previous "Task N-1", is the n value calculated by equations in ***Table 2. 1***.

Table 2. 1 Conditional probability of success and failure on Task N (Swain, 1964)

Level of Dependence	Success Equations	Failure Equations
ZD	$P[S_N \mid S_{N-1} \mid ZD] = n$	$P[F_N \mid F_{N-1} \mid ZD] = n$
LD	$P[S_N \mid S_{N-1} \mid LD] = \frac{1+19n}{20}$	$P[F_N \mid F_{N-1} \mid LD] = \frac{1+19n}{20}$
MD	$P[S_N \mid S_{N-1} \mid MD] = \frac{1+6n}{7}$	$P[F_N \mid F_{N-1} \mid MD] = \frac{1+6n}{7}$
HD	$P[S_N \mid S_{N-1} \mid HD] = \frac{1+n}{2}$	$P[F_N \mid F_{N-1} \mid HD] = \frac{1+n}{2}$
CD	$P[S_N \mid S_{N-1} \mid CD] = 1$	$P[F_N \mid F_{N-1} \mid CD] = 1$

3) Recovery factors

THERP is the technique that emphasises unique error expression and error recovery through the Human Reliability Analysis Event Tree (HRAET). Despite the THERP's advantages and characteristics, the approach is overly thorough and time-consuming due to the amount of work required to create HEP values using THERP analysis. Furthermore, it gives the operator a limited nominal failure probability in nuclear control rooms. To apply this method to the new systems, it is subsequently necessary to collect the human behaviour data required for human reliability studies (Swain, 1964, Kirwan, 1994). The THERP is utilised as a database of human errors in maritime research ((Zhang et al., 2020b, Zhang et al., 2020a, Martins and Maturana, 2010), despite limited PSFs and human activities pose barriers to implementation in a broader field.

2.5.2 CREAM

The Cognitive reliability and error analysis method (CREAM) was first proposed by Hollnagel (1998), which was initially developed for nuclear power plant applications (He et al. (2008); Lee et al. (2011a); Tang et al. (2014)). This method was adopted by the National Aeronautics and Space Administration (NASA) in the early 1990s to predict human error (Calhoun et al., 2014). The CREAM is a commonly used HRA method and has a convenient structure to integrate other techniques for improvement. For human error quantification, experts are recruited and asked to assess the context, referred to as common performance conditions (CPCs). Human performance cannot be predicted in the absence of context. So it is fair to examine a technique in which the "error probability" may be estimated directly from the features of components that influence human performance since the context may serve as an "error forcing condition" that causes the failure to occur (Fujita and Hollnagel, 2004). Afterwards, there are two distinct CREAM ways to choose. The CREAM extended technique tries to generate particular action failure probabilities. In contrast, the basic method does not consider individual human actions when estimating the action failure probability. Instead, it relies solely on a context evaluation to make the human error mode prediction. A more in-depth investigation of the screening process using the human error probability acquired through the CREAM basic method or examining specific event sequences may be accomplished by employing the CREAM extended approach (Ahn and Kurt, 2020). The CREAM has two characteristics called CPCs and COCOM, as below.

1) Common Performance Conditions (CPC)

As shown below, the original CREAM divides the effect of human reliability into nine components, which are referred to as CPCs (Common Performance Conditions).

(i) Adequacy of organisation,

(ii) Working conditions,

(iii) Adequacy of a man-machine interface (MMI) and operational support,

(iv) Availability of procedures and plans,

(v) Number of simultaneous goals,

(vi) Available time,

(vii) Time of day,

(viii) Adequacy of training and experience, and

(viii) Crew collaboration quality.

Three distinct CPC levels correlate to three different impacts on human behaviour reliability: negative (reduce reliability), neutral (neutral), and positive (raise reliability) (Hollnagel, 1998).

2)Contextual Control Mode (COCOM)

According to the status of human cognition and action, there are four typical control modes that the CPCs identify, namely "Scrambled," "Opportunistic," "Tactical," and "Strategic." The CPCs determine these four control modes based on the type of human cognition and activity. Furthermore, as indicated in ***Table 2. 2***, the control modes are linked to distinct failure probability ranges representing human action failure probabilities (Hollnagel, 1998). As a result, the control modes may be used to characterise the performance of a team or group of individuals in the same way that they can characterise the performance of an individual. The subsequent phases of CREAM are described in depth in Chapters 6 and 7, with case studies to illustrate each step.

Table 2. 2 Control *mode and action failure probability (Hollnagel, 1998)*

Control mode	Action failure probability
Strategic	0.5E-5< p < 1.0E-2
Tactical	1.0E-3< p < 1.0E-1
Opportunistic	1.0E-2< p < 0.5E-0
Scrambled	1.0E-1< p < 1.0E-0

2.5.3 SPAR-H

The Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method was developed to estimate the human error probabilities associated with operator and crew actions and decisions in response to initiating events at commercial U.S. nuclear power plants by Blackman et al. (2008). The SPAR-H technique simplifies the computation of HEP rates by beginning with predefined nominal error rates for cognitive versus action-oriented tasks and multiplying those nominal error rates with performance shaping factors (Blackman et al., 2008). The SPAR-H categorises Human Failure Events (HFE) as either Diagnosis tasks, Action tasks, or a combination of Diagnosis and Action activities (Whaley et al., 2011). Once HFEs are categorised, analysts identify factors that affect human performance positively and negatively to support qualitative evaluation. This process can be supported by reviewing SPAR-H performance shaping factors. These factors include eight PSFs: time available, stressors, experience and training, complexity, ergonomics including human-mechanical interfaces, procedures, fitness for duty, and work processes. When the PSFs level is specified, the final HEP is the product of the nominal HEP and the composite multipliers of PSF, with the following equations:

HEP = Diagnosis Error + Execution Error (2-2)

Diagnosis Error = Nominal Diagnosis Error x Composite Multipliers of PSFs (2-3)

Execution Error = Nominal Execution Error x Composite Multipliers of PSFs (2-4)

When there are more than three negative PSFs, human error probability needs to be adjusted using equation (2-5).

$$\text{Adjusted HEP} = \frac{0.01*\prod multipliers\ of\ PSFs}{0.01*(\prod multipliers\ of\ \ PSFs-1)+1} + \frac{0.001*\prod multipliers\ of\ PSFs}{0.001*(\prod multipliers\ of\ \ PSFs-1)+1} \qquad (2\text{-}5)$$

However, one of the method's drawbacks is the SPAR-H designed for the nuclear industry. As a result, it may not be fully applicable to all maritime operations. As a result, further study is required to design and validate customised PSFs for the maritime industry's more diversified working circumstances.

2.5.4 SLIM

The Success Likelihood Index Method (SLIM) is a technique used in the HRA field to evaluate the probability of a human error throughout the completion of a specific task. SLIM provides a set of models for the factors that influence human error during commonly occurring activities, including alarm response, actions, checking, information retrieval, and communication. The SLIM is a decision-analytic approach to quantifying PSFs using expert judgement, and factors related to an individual, environment, or task are likely to positively or negatively impact human performance. These factors are used to derive a Success Likelihood Index (SLI), a form of preference index corrected for existing data to derive a final Human Error Probability (HEP). The PSFs that must vitally be considered are chosen by experts and are those factors that are regarded as most significant concerning the context in question. Performance Shaping Factors (PSF), which have a significant impact on human performance, can be quantified in SLIM and converted to a preference index form (Akyuz, 2016), allowing for the quantification of external factors affecting human performance that are quantitatively reflected in the form of human error probability. The SLIM and SLIM-MAUD are particularly beneficial for discovering which condition of PSFs minimises the HEP most efficiently. The SLIM consists of six steps: 1) task analysis and scenario development; 2) derivation of the PSF; 3) rating of the PSF; 4) weighting of the PSF; 5) computation of the SLI; and 6) conversion of the SLI to HEP.

1) Task analysis and scenario definition

In this phase, task analysis aims to simply subdivide functions into tasks, tasks into subtasks, and subtasks into human behaviours. A task analysis explains the processes involved in an activity, offering a technique for organising the data gathered about the task methodically (Bye et al., 2017). Whereas the scenario establishes the scope and limitations of the analysis and serves as the foundational material for later qualitative and quantitative studies (Bye et al., 2017). Through the process, this stage is concerned with describing the context in which specific activities are performed. The primary goal of scenario development is to provide a complete description of the event sequence to identify potential human errors better and explain the operational context.

2) PSF derivation

This step aims to identify the PSFs that have the most influence on the tasks rather than to elicit all conceivable influencing variables. A panel of experts is tasked with identifying a set of PSFs that are appropriate for the task at hand within the context of the more extensive system, and then selecting some of the most significant PSFs in light of the scenario's conditions.

3) PSF rating

This is a stage in which the level of each PSF is determined. Experts assigned values between 1 and 9 on a linear scale to the identified endpoints of each PSF. Based on their judgment, the expert is needed to offer a rating to each task between the two endpoints that appropriately reflects the conditions present during the task in question. It is preferable to analyse each aspect so that the judgments made are not influenced by other elements that might sway opinion.

4) PSF weighting

This phase is used to determine the relative significance of each PSF, i.e., the amount of influence each PSF has on the success of a task because not all PSFs have the same effect on human performance. When experts perform this function and their opinions are altered, the weighting of each PSF's influence on task success may be inferred. This stage should be completed iteratively to improve the accuracy of the outcome.

5) SLI calculation

After the rating (R_i) and relative importance (W_i) of PSF are determined, the Success likelihood Index (SLI) for each task is derived by the equation given below.

$$SLI = \sum_{i=1}^{9} \mathrm{Ri} \times \mathrm{Wi} \qquad (2\text{-}6)$$

6) Conversion of SLI into HEP

The SLI value is converted into the HEP value as below.

Log of Probability of Success = a*SLI + b (2-7)

where a and b are constant (Embrey et al., 1984).

The SLIM is a highly practical and straightforward approach for calculating human error in situations when obtaining human error data is challenging (Park and in Lee, 2008). While the SLIM can generate data on human error, the numbers at both ends must eventually be derived from past human data. Due to these limits, the HRA data scarcity issue cannot be resolved entirely. Additionally, even after implementing the SLIM, the issue of expert dependency persists. Chapter 9 will get into further detail on the SLIM.

2.5.5 HEART

The Human Error Assessment and Reduction Technique (HEART) is an HRA technique developed to identify the contribution to human performance and the likelihood of error in a systematic and repeatable way (Bell and Williams, 2016). This method is based on the general principle that every task has a nominal failure probability. Various levels of Error Producing Conditions (EPCs) influence each of these tasks, affecting human performance in systems operations. The method provides users with human reliability data that can be modified to be specific to their risks. The HEART is a relatively quick and straightforward method that is applicable to any industry where human reliability is essential. The steps of obtaining human error based on the HEART method can be summarised as follows.

Step. 1 Determine HEART general task type

Tasks are classified into eight categories based on the general HEART task type known as the GTT. Each task type assigns a nominal probability of human error to a particular type of human action.

Step. 2 EPCs selection

When a specific type of HEART task is set up, elements called the EPCs must be found. Error Producing Conditions (EPCs) are variables that are linked to situations or tasks that could have an impact on performance. Then, multipliers are calculated by examining the number of Error

Producing Conditions (EPCs) present in the circumstance. HEART includes a multiplier for the thirty-eight (38) kinds of EPCs.

Step 3 Assess the proportion of EPC's effect

This step is to determine how much the selected EPC affects GTT. This step is similar to evaluating the context. This process estimates the assessed proportion of effect, and the assessed impact per EPC is computed below.

Assessed impact (AI_i) = (($Multiplier_i$ −1) x $Assessed\ proportion\ of\ effect_i$) +1 (2-8)

Then, based on HEART, human error probabilities for each task can be calculated as below.

HEP = Nominal human failure probability × AI_1 x AI_2…AI_i (2-9)

The HEART is not a limited method to the nuclear field despite the occasional opposition, even though it came from the nuclear industry. However, there are still limitations to applying the HEART to new technologies system, although it is a very well-structured method. For example, with a limited number of GTT data; only eight types of nominal human error probability, EPC does not reflect the present technology context. Thus, it is necessary to continue collecting data using various methods. For instance, more accident data should be collected and efficiently shared using a unified dataset platform. In addition, state-of-the-art technologies such as simulators should be used to verify the magnitude of the impact of new types of factors on human performance. Nonetheless, there is little doubt that HEART is one of the most comprehensive and well-established methodologies of human reliability assessment now available.

2.6 Systematic approaches for human reliability

Modern process systems are confronted with new safety issues due to new technologies. Maritime operating systems have grown increasingly software-intensive, consisting of not just hardware components but also logic control devices, software, and a rising number of sensors (Sultana et al., 2019). The increasing speed of technology advancement and the evolution of more complicated connections between people and automation are eroding the capabilities of current accident models and safety engineering technologies, necessitating the development of novel alternatives (Leveson, 2004). It is difficult to evaluate the reliability of complex systems, that is, systems in which humans, machines, and software interact, just by observing human errors or equipment failures. Therefore, a comprehensive analytic technique is necessary for complex systems (Kirwan, 1994). In this context, the Functional Resonance Analysis Method (FRAM) by Hollnagel (2017a) and System Theoretic Process Analysis (STPA) by Leveson and Thomas (2018) present novel methodologies relevant to these socio-technical and complex systems.

2.6.1 FRAM

The Functional Resonance Analysis Method (FRAM) is a Resilience Engineering technique that begins by describing a system's typical functions and looks for ways to improve the system's capacity to adapt, monitor, learn, and predict (Hollnagel, 2017a). The objective of resilience engineering is to guarantee that an organisation can function efficiently under normal operating conditions and to ensure that routine work is completed correctly (Hollnagel, 2017b). As a result, the FRAM attempted to explain the system in terms of work-as-done rather than anticipated. A functional model can be described in six distinct ways, as seen in ***Figure 2. 7*** and an example of the FRAM model for firefighting on a ship as illustrated in ***Figure 2. 8***. The FRAM was also used in the following research relating to maritime safety. Praetorius (2014) utilised the FRAM to comprehend how variability in its functional units affects the VTS service's overall system performance. Badokhon (2018) proposed a model for incorporating resilience engineering principles with ship management systems in his PhD thesis to improve the navigational bridge. Lee and Chung (2018) attempted to quantify the influence of variability in human-system interactions using the FRAM for maritime incidents. Salihoglu and Beşikçi (2021) applied the FRAM for qualitative risk analysis to the Prestige oil spill accident.

Using the FRAM technique, Lee et al. (2020) investigated human collaboration in maritime operations. Qiao et al. (2022) integrated FRAM and a BN to conduct a resilience evaluation of maritime liquid cargo emergency response. However, it does not address how to apply the results of FRAM analysis in practice or how to replace the existing risk assessment framework quantitatively. Because the FRAM approach emphasises variability of performance rather than probability due to the uncertainty surrounding the human and organisational contribution to system failure. In this regard, Praetorius et al. (2017) applied FRAM to Formal Safety Assessment (FSA), a risk assessment methodology that is widely employed in the modern maritime sector. The findings show that FRAM may be considered as a complement to traditional risk assessment approaches such as FTAs, but probably not as a standalone method suitable for the FSA. As a result, a more precise framework for measuring human and overall system reliability should be established to replace present methodologies.

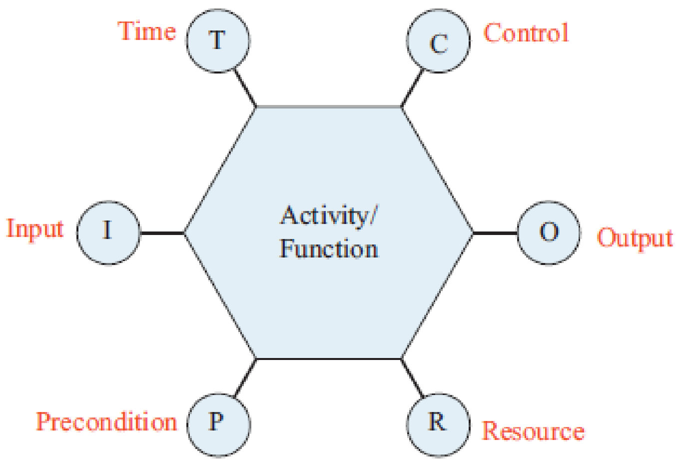

Figure 2. 7 Functional model in FRAM approach(Hollnagel, 2017a)

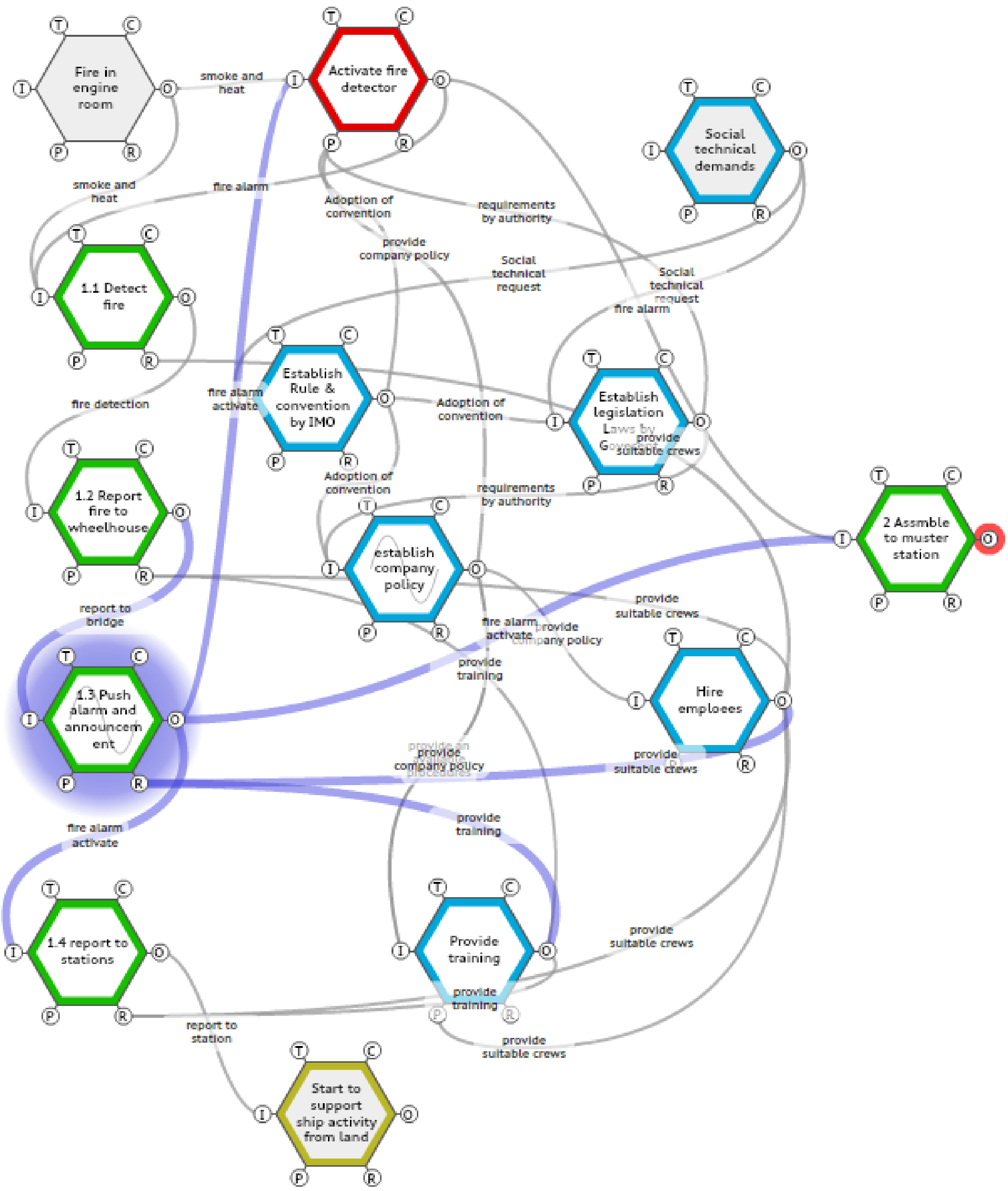

Figure 2. 8 Example of instantiation of the FRAM model for fire-fighting on ship

2.6.2 STAMP

The System-Theoretic Accident Model and Processes (STAMP) is the name of a novel model of accident causation based on systems theory; safety is viewed as a dynamic control challenge rather than a failure avoidance problem. Nonetheless, the STAMP is not an analytic technique but the theoretical basis for analysis (Leveson and Thomas, 2018). Today, the two STAMP-based techniques that are most extensively used are System Theoretic Process Analysis (STPA) and Causal Analysis based on Systems Theory (CAST). The STPA is a proactive analytic technique that identifies probable accident causes during development to minimise or mitigate risks. The CAST is a way of conducting retrospective accident analyses to determine the underlying causative elements. The STPA assumes that accidents can also be caused by hazardous interactions between system components, none of which may have failed in addition to component failures. The STAMP-based methods have been applied to maritime research, including an autonomous ship and a variety of activities to identify hazards and safety requirements and examine the causes of an accident. For example, Chaal et al. (2020) developed a framework for simulating an autonomous ship's STPA hierarchical control system. Dghaym et al. (2021) used STPA to establish safety and security needs for an autonomous maritime system. Furthermore, Rokseth et al. (2017) examined the viability of employing a systematic approach to dynamically positioned systems, while Gil et al. (2019) assessed control actions for ship collision avoidance using an STPA-based model.

Even though the STPA finds additional hazards that conventional risk analysis methods have missed and accurately represents the impacts of human and organisational elements, it requires supplementation to fully utilise the STPA approach to risk identification, quantification, and reduction. This is not only a problem of quantification of error. As previously stated, humans play a critical role in safety-related systems such as maritime operations. The STPA evaluates humans from a controller perspective. On the other hand, human decision-making processes and behaviour are far more complicated than the process model of the software controller, and human performance is predominantly influenced by a factor called Performance Shaping Factors (PSF). Thus, while assessing the total reliability of the maritime system, an integrated approach should be used, including human reliability assessment (HRA). The case study for the integrated model will be presented in Chapter 9.

2.7 Past HRA studies in the maritime

Over the decades, significant effort has been made to understand the mechanism of human error and to prevent maritime accidents caused by humans by utilising various human reliability assessment (HRA) techniques, such as Success Likelihood Index Method (SLIM), Human Error Assessment and Reduction Technique (HEART), Technique of Human Error Rate Prediction (THERP), Human Factors Analysis and Classification System (HFACS) and Cognitive reliability and error analysis method (CREAM). By examining and analysing past data, investigators can methodically identify active and latent flaws that result in an accident (Kirwan and Ainsworth, 1992). However, accurate historical data collection is challenging since human and organisational factors influencing accident development and emergency response are not reported frequently (Schröder-Hinrichs et al., 2011). As a result, Human Reliability Assessments (HRAs) are a source of concern for safety engineers and risk assessment analysts due to fundamental limitations such as insufficient data, methodological limitations due to analyst subjectivity and expert judgement, and restrictions due to uncertainty about actual human behaviour under accident conditions (Konstandinidou et al., 2006). Meanwhile, HRAs are composed of three critical components: identification, quantification, and reduction of human error (Kirwan, 1994). Therefore, different studies have been conducted to address the research gap between the various elements of HRAs. The following paragraph summarises previous maritime studies performed using the HRA method.

The Human Factors Analysis and Classification System (HFACS) was developed by Shappell and Wiegmann (2000) and provides a powerful tool for assisting in the investigative process for accident records and focusing on training and preventative efforts. The HFACS describes human error at the same level as Reason's Swiss Cheese Model (Reason, 1990) at four levels of failure: organisational influences, unsafe supervision, a precondition for unsafe acts, and unsafe acts. In this approach, human error is not regarded as the cause of accidents but rather a symptom of an organisational problem. It is taxonomic to facilitate an understanding of human behaviour. Several academics suggested combining the HFACS with a Fuzzy Analytical Hierarchy Process (FAHP) or Fault Tree Analysis (FTA) to get quantifiable results. For example, Celik and Cebi (2009) developed an analytical HFACS based on the FAHP idea to determine the influence of human error in bulk carrier boiler explosions. This study aims to establish an analytical foundation and collective decision-making capability to conduct a quantitative analysis of maritime accidents. Akyuz and Celik (2014) used a hybrid model of HFACS and cognitive mapping to detect the spread of human error in maritime accident investigations. Zhang et al. (2019) introduced a modified model of the HFACS for collision accidents between a ship and an icebreaker. Then, the FTA model was utilised to analyse the fundamental collision risk factors according to the statistical analysis of accident reports and experts' judgment based on the HFACS-SIBCI model. Collision risk factors during icebreaker assistance were identified and classified under the initial HFACS framework. However, previous research indicates that maritime taxonomies are still in their infancy, so HFACS will fall short of adequately addressing maritime concerns. For instance, Salmon et al. (2012) highlighted that one of the primary challenges associated with applying HFACS outside of aviation is that it was built exclusively for aviation. In other words, numerous error and failure mechanisms are unique to the aviation sector.

The Success Likelihood Index Method (SLIM) is used in the HRA field to evaluate the probability of a human error throughout the completion of a specific task, as described in section 2.5.4. The SLIM has also been frequently used in the maritime sector because of the benefits described above. For example, Abbassi et al. (2015) demonstrated a combined HRA technique for offshore condensate pump maintenance work by merging the SLIM and the Technique of Human Error Rate Prediction (THERP). This strategy is built on several data collectors to exploit THERP's quantitative data and produce human error data in SLIM without existing data. Akyuz (2016) applied the concept of the SLIM for estimating HEP when conducting the abandon-ship procedures. The Fuzzy sets were used to improve the reliability of the analysis against the vagueness of expert judgments and the arbitrary measure of performance shaping factors (PSFs). Based on the SLIM, Islam et al. (2016) determined the HEPs related to marine engine maintenance tasks. In another study, Islam et al. (2017b) developed a monograph to assess the likelihood of human error in maritime operations applicable to instant decision-making. It was identified that with the SLIM method, it is possible to estimate general HEPs in each context and HEPs in specific activities by adding PSFs, such as training, experience, fatigue level of a seafarer, etc. However, SLIM is overly relying on expert judgment, which makes the analysis results highly subjective and less reliable because the scope of PSFs is limited to certain contexts rather than fully reflective to every aspect that affects human performance. They are weak in dealing with social and organisational aspects.

The Human Error Assessment and Reduction Technique (HEART) is an HRA method that has also been frequently used in the maritime sector because it is a well-established HRA method, as discussed in section 2.5.5. For example, Noroozi et al. (2014) applied HEART analysis to human error during maritime maintenance operations. Akyuz and Celik (2016) also introduced the HEART application, combining the Analytic Hierarchy Process (AHP) to the case of a cargo loading operation in an oil/chemical tanker ship for human error probabilities estimation. Islam et al. (2017a) developed an operational-specific methodology based on HEART to capture unique features of maritime environments and operations and apply it to the maintenance procedures of a marine engine exhaust turbocharger and a condensate pump on offshore oil and gas facilities. On the other hand, Akyuz and Celik (2016) applied the HEART in a combination of AHP to predict human errors associated with cargo operation on oil/chemical tankers. The HEART is similar to the SLIM but provides nominal probabilities for generic HEART tasks. After that, the overall HEPs are adjusted by evaluating Error Producing Conditions (EPCs) and the proportion of effect defined by experts' judgment. As a result, like the SLIM, the multiplier values are highly reliant on experts' knowledge, which leaves uncertainties in analysis results.

According to the past research presented above, it can be concluded that the first generation HRA methods have relied on context assessment to estimate HEP and/or to determine performance shaping factors that may cause human errors or misbehaviours against certain features of the maritime tasks. However, those tools are less considerate of organisational factors and their interaction among PSFs. To remedy the weakness of the first generation methods, the cognitive reliability and error analysis method (CREAM) has been introduced as the second HRA generation, where the individual events and their success or failures are further detailed and examined. The CREAM provides a framework of the subjective HEP estimation from expert judgement by evaluating PSFs in the basic method and provides a nominal probability for each subtask if the subtask is converted to one of the cognitive activities. This means the CREAM makes it possible to estimate overall HEP by evaluating context with PSFs. At the same time, CREAM provides nominal probabilities for cognitive activities. This makes it possible to generate more reliable data, especially useful when there is an unavailability of past data.

The Cognitive Reliability and Error Analysis Method (CREAM) was first developed by Hollnagel (1998) to predict human performance reliability. The human error probability can be determined directly from a characterisation of the context based on a description of the specific circumstances or conditions (Fujita and Hollnagel (2004)). Since the introduction of the initial concept of the CREAM, numerous follow-up studies have been conducted in different disciplines to achieve highly advanced CREAM methods through which HEPs could be combined in different ways, such as giving customised changes to reflect characteristics of the specific industry and its application to critical operations. Yang et al. (2013) proposed a modified CREAM to facilitate human reliability quantification in marine engineering by incorporating Fuzzy evidential reasoning and a Bayesian network based on inference logic. They extend the traditional CREAM method to a Fuzzy environment to quantify human failure probabilities by integrating Bayesian reasoning to model the dependency among CPCs. The multiple-input multiple-output rule concept, together with evidential reasoning, estimates human failure probabilities as reasonable in the way of being sensitive to the minor changes of Fuzzy input. It also makes it possible to realise the instant calculation of human failure probabilities in specific task analysis onboard ships. The developed method was demonstrated by an illustrative example of an oil tanker's Cargo Oil Pumps (COPs) shutdown scenario. Ung and Shen (2011) proposed a systematic procedure to compute CREAM's probability of operator action failure. Then, in a further study, Ung (2015) developed a weighted Fuzzy CREAM method. The features of the model mentioned above include; the consideration of the weight of each CPC, refinement of the logicality between the CPCs and Contextual Control Modes (COCOM) and the deliberations of helpful information from each input for the oil tanker's COPs shutdown scenario same with the scenario of Yang et al. (2013). Furthermore, Zhou et al. (2017a) adopted the eight customised CPCs to better capture the essential aspects of the work situations and conditions onboard tankers with the weighting of the CPCs by employing the Fuzzy Analytical Hierarchy Process (FAHP). Lee et al. (2011a) suggested a customised CPC called Cognitive Speaking Process (CSP), focusing on communication errors in a nuclear plant. Some studies illustrated a risk assessment combining the CREAM method. For example, Zhou et al. (2017b) utilised the CREAM method with a modified fault tree model for LNG spill accidents during LNG carriers' handling operations for risk assessment. Ung (2019) demonstrated risk assessments of human error contribution to oil tanker collision by using the Fault Tree Analysis (FTA) structure under which a modified Fuzzy Bayesian network is also based on Cognitive Reliability Error Analysis Method (CREAM). Even though newly developed CREAM methods can be considered more reliable and sensitive quantification models, most advanced and modified CREAM methods focus on the CREAM basic method to predict overall HEPs by evaluating contexts. Hence, they would fail to utilise the extended CREAM method to predict individual cognitive failure probability for each task in operating procedures.

Meanwhile, a simplified CREAM method introduced by He et al. (2008) provided a different view of the CREAM basic and extended method. Akyuz (2015) and Akyuz and Celik (2015) analysed the critical maritime operating procedures by adopting both simplified CREAM basic and extended methods. Xi et al. (2017) introduced a modified CREAM methodology utilising an Evidential Reasoning (ER) approach and a Decision-Making Trial and Evaluation Laboratory (DEMATEL) technique to make human error probability quantification in CREAM rational, which applies to the CREAM basic and extended method. A simplified CREAM method is an easily accessible process to obtain the numeric results, but numerous assumptions were inevitably made to estimate the uncertainties posed by the over-simplification idea. For instance, two distinct situations with the same level of negative are deemed to have the same failure probability. This is because the outcome does not accurately reflect the features of individual events. As demonstrated in ***Table 2. 3***, a variety of human dependability assessment techniques are employed in the maritime industry, from emergency response to vital operations and maintenance. The reviewed maritime human reliability assessment methods focus on human error quantification because HRA is dealt with within the maritime probabilistic risk assessment framework. The contribution of past human reliability assessments in the maritime sector mainly dealt with uncertainties using expert judgment, assignment of nominal failure probability for specific tasks, and selection of performance shaping factors that affect human performance.

Table 2. 3 Maritime HRA studies

Authors	**Maritime case studies**	**Applied HRA methods**
Yang et al. (2013)	Oil tanker’s Cargo Oil Pumps shutdown scenario	CREAM incorporated with Fuzzy evidential reasoning and Bayesian network based on inference logic
Noroozi et al. (2014)	Maritime maintenance operations	HEART
Ung (2015)	Oil tanker’s Cargo Oil Pumps shutdown scenario	Weighted Fuzzy CREAM
Akyuz and Celik (2015)	The cargo loading process of the LPG tanker	Quantified CREAM utilising a context influence index
Abbassi et al. (2015)	An offshore condensate pump maintenance task	Integrating the SLIM with the Technique of Human Error Rate Prediction (THERP)
Akyuz (2016)	The abandon-ship procedures	Fuzzy SLIM
Islam et al. (2016)	Marine engine maintenance tasks	SLIM

Akyuz et al. (2016)	the maintenance procedures of a marine engine exhaust turbocharger and a condensate pump fitted to offshore oil and gas facilities	HEART in a combination of AHP
Wu et al. (2017)	Ship capsizing accident	CREAM incorporated with Fuzzy evidential reasoning
Xi et al. (2017)	The collision avoidance of a particular scenario in Shanghai coastal waters	Modified CREAM based on an Evidential Reasoning (ER) approach and a Decision-Making Trial and Evaluation Laboratory (DEMATEL) technique
Zhou et al. (2017a)	The general seafarers' human reliability when performing tasks under the operation circumstance in tanker shipping	Quantified CREAM incorporated with Fuzzy analytical hierarchy process (FAHP) for the weighting of the CPCs
Zhou et al. (2017b)	LNG carrier spill accidents	Incorporating CREAM and MCS into fault tree analysis
Yang et al. (2019)	Drilling rig crew's actions in monitoring the Macondo well and managing the well control event on April 20 2010	CREAM is based on Evidential reasoning for eliciting Bayesian subjective probabilities
Zhang, Zhang et al. (2019)	Accident analysis for the collision accidents between a ship and an icebreaker	HFACS with Fault tree analysis
Ung (2019)	Oil tanker collision	Fault tree analysis and modified Fuzzy Bayesian Network-based CREAM
Ahn and Kurt (2020)	Engine room fire drill	Fuzzy CREAM
Liu et al. (2021)	Maritime autonomous surface ship	SLIM under an interval type-2 Fuzzy sets approach
Uflaz et al. (2022)	Ship navigation	Fuzzy AHP-based shipboard human reliability analysis
Ahn et al. (2022)	Rescue boat drill	Integration of SPAR-H into a risk model

However, human error quantification techniques rely either on expert judgment or on a combination of data and psychology-based models, which assess the main impact of human performance (Kirwan, 1994). Researchers have used techniques like Fuzzy logic, Bayesian networks, Evidential inference, Event tree, Fault tree, and other forms of integration to turn complex situations where people are likely to make mistakes into quantified human error probabilities. The Fuzzy theory and Bayesian network are commonly integrated into HRA techniques to develop more advanced human reliability methods. For example, Fuzzy logic is utilised to convert qualitative data to quantitative data and opinion aggregation from multiple groups with combined HRA techniques. At the same time, Bayesian networks are applied to consider dependency and weighting among PSFs. These techniques enhance the consistency of research, minimise subjectivity and ambiguity during expert judgment, and provide instant calculations of human error probability.

Fuzzy logic has been successfully applied to a wide range of maritime safety and risk topics. For example, Balmat et al. (2011) presented a Fuzzy approach to evaluate the maritime risk assessment to pollution prevention on the open sea, while Wu et al. (2019) utilised Fuzzy Multiple Attribute Decision Making for a ship-bridge collision alert system. Furthermore, in numerous studies on human reliability analysis, Fuzzy logic has also been utilised to improve reliability and reduce uncertainty in results.

There have also been attempts to develop models that could directly estimate overall HEPs using BNs. Islam et al. (2018) introduced a BN model to estimate HEP using expert groups' priority probability and Conditional Probability Table (CPT). It determined the impact of internal and external factors on human performance with a case study for ship maintenance activities. Unlike the HRA studies mentioned above, Vagias (2010) investigated specific aspects of human fatigue. The Bayesian Network (BN) was utilised to predict fatigue prevalence and its importance, given the workload, environment, and ergonomic factors before the accident. This study also provides comprehensive information about human factors and human error. The BN model provides flexible HEPs that could be assigned with new input variables. As such, it made possible to predict HEPs across various maritime scenarios dynamically. Despite its effectiveness on HEPs, the BN models are subject to uniformed contexts, thereby the same level of PSFs, against disparate activities. Moreover, the direct inference logic model is hard to figure out the significant difference among subtasks under a similar situation without considering different tasks because contributing factors do not fully address the characteristic of the different levels of tasks. Furthermore, de Maya et al. (2019b) proposed the MALFCM approach incorporated with BNs based on the concept and principles of Fuzzy Cognitive Maps (FCMs) to represent the interrelations amongst accident contributor factors. Although this database-driven research has led to successful results, the applicable range of the database is far limited to some specific cases rather than general ones, which can be highlighted as a weakness.

Once human errors are examined, the next step is a human error representation that uses modelling to carry out a risk assessment to reduce error. Some studies have demonstrated a risk assessment combining the human reliability assessment methods. For example, Zhou et al. (2017b) utilised the CREAM method with a modified fault tree model for LNG spill accidents during LNG carriers' handling operations for risk assessment. Ung (2019) applied Fault-tree analysis where a modified Fuzzy Bayesian network-based CREAM was applied to a risk assessment of human error contribution in oil tanker collisions. Human reliability assessment is expected to evaluate the degree to which humans contribute negatively or positively to the system by considering human roles in the entire system and then integrating them into the overall risk picture. Meanwhile, less attention has been paid to error representation for system elements such as human and machine interactions and human recovery actions.

2.8 Challenges in HRA application

Several additional features of the problems revealed in the application of HRA from a prior literature review are outlined below. It addresses both the problem of HRA research and the obstacles associated with HRA implementation in the maritime sector.

1) Human error data-driven perspective

One of the main problems with HRA research is data availability. There is a shortage of historical data on human errors particularly in the maritime sector because these data have not been systematically accumulated and managed. Other than historical data, additional sources of information may include simulator data; however, such simulators are usually primarily utilised for operator re-authentication reasons, such as training or in the US nuclear field. In addition, human error literature has a similar problem in this field as it is usually highly controlled. One or two independent variables are frequently viewed (unlike industries with many changes and interactions). Good motivational topics are used for a short period. It is not easy to generalise considering the complex industrial and multi-person situations in these studies, and it is a truly questionable strategy. Therefore, most HRA studies rely heavily on subjective expert judgment when predicting human errors in a specific context. This leads to a problem of uncertainty in results in terms of human errors and ratings of performance shaping factors that affect human performance.

The following causes account for the inaccessibility of such data:

- Difficulties involved in estimating the number of opportunities for error in realistically complex tasks (the so-called denominator problem).
- Confidentiality.
- An unwillingness to publish data on poor performance.
- A lack of awareness of why it would be useful to collect such data in the first place, and hence a lack of financial resources for such data collection.

There are several other potential reasons too, and one can refer to Kirwan et al. (1990) for further details.

2) Human error quantification perspective

Human error is linked directly or indirectly to several elements referred to as performance shaping factors (PSFs). PSF is the most often utilised as a means of assessing human errors. These PSFs are characteristics of human behaviour and situations that might influence human performance. They are frequently used to calculate human error probability (HEPs) and identify contributors to human performance. To predict the reliability of human performance, a contextual description must be provided, as predictions of what could occur in a certain condition or situation must be based on a description of the specific circumstances (Hollnagel (1998); Fujita and Hollnagel (2004)). It is reasonable, then, that human error probability can be affected by characterising the context. Given the unique operating conditions of the shipping industry, in which seafarers face many dangerous situations, proper PSF evaluation is vital in estimating human error. However, the limitation of the application of the maritime industry is that these PSFs were not developed based on the unique characteristics of the maritime sector. Once PSFs are identified, an evaluation is conducted based on the expert's judgment. However, uncertainty and inconsistency are inevitable due to the lack of objective data mentioned during the human error evaluation process. In addition, there is no systematic model that enables consideration of dependence due to interactions between tasks, people, or system elements. Since measuring human error is the most crucial aspect of HRA analysis, it will be discussed in further detail in Chapters 6, 7, 8, and 9 utilising the various indicators for determining it.

3) Human error identification in a complex system

New technologies can help improve maritime safety, but we should also consider identifying human roles in complex systems such as autonomous software-intensive systems. This is because many human roles are difficult to find only with traditional approaches such as hierarchical work analysis commonly used in the existing HRA of future complex systems. As a result, finding a human role will be more challenging. Furthermore, by introducing autonomous or software control systems in the maritime industry, human functions are increasingly required for cognitive-intensive behaviour in addition to existing labour-intensive behaviour. Therefore, it is necessary to examine what improvements have been made to apply the existing HRA to the changed or added human role.

4) human error representation perspective

HRA techniques mainly focus on quantifying human errors but do not deal well with evaluating human reliability. Even though the chance of error is identical, the total reliability depends on the design of the system and the interdependence between each task; hence, no evaluation of human reliability, such as deducting the probability of error from one, should be produced. The model to integrate human error into the PRA framework is needed for applying probabilistic risk assessment. The matter of how HRA is integrated into risk assessment is dealt with in more detail in Chapters 8 & 9.

5) different needs for HRA

Each HRA technique is created for a specific purpose. Therefore, although many methods do not completely cover the qualitative and quantitative stages, they should not be considered deficient (Boring, 2010). Nevertheless, choosing the HRA method that meets study needs takes much effort. However, suppose an optimised method is provided to respond to the needs of each study or analysis. In that case, HRA research and application will be very active. However, it is difficult to find past studies on this part, and in this paper, Chapter 5 will provide guides and discussion on this part.

2.9 Chapter summary

This chapter analysed the evolution of the general HRA method and previous studies on HRA in the maritime industry. In addition, study gaps and shortcomings in earlier research cases were found. The aim and objectives of this research study will be discussed in Chapter 3.

3 Research aims and objectives

3.1 Chapter overview

For the successful development of this study, it is important to identify the problems in Maritime HRA to be solved within this study, along with the objectives to be used as a milestone to achieve the research aim of this study. Therefore, the motivation for this study and research questions will be presented in Section 3.2. In addition, Section 3.3 outlines this study's overall aims and objectives.

3.2 Motivation and research questions

As previously discussed through literature reviews, various HRA techniques are applied to the maritime sector to improve safety. The main contribution of past human reliability assessment in the maritime sector was dealing with uncertainties using expert judgment, assignment of nominal failure probability for specific tasks, and derivation of performance shaping factors that affect human performance. Nevertheless, there are remaining questions to be solved in applying HRA in the Maritime industry. First, with the introduction of changing technologies, the interface between humans and machines rapidly evolves into computer-based communication, and more human cognitive activities are required. In the end, this means a change in the working context that profoundly influences human performance. Therefore, examining the changing working environment and process in more detail and the newly occurring factors affecting humans is necessary. Additionally, the previously created PSFs are not based on human activities in the maritime sector but on task characteristics and settings in the nuclear industry. As a result, it may not be completely applicable to all maritime operations. Another difficulty is effectively incorporating the entire risk of human error into the probabilistic risk assessment framework. However, there was less interest in and research on error representation than human error quantification.

Although the points mentioned above focused on improving the HRA method, the view that it is an appropriate choice of HRA should not be ignored. For instance, any sophisticated HRA approach might result in under- or over-analysis and is inefficient if a uniform HRA analysis method is applied without considering the unique characteristics and analysis goal of the operation, system, or facility under examination. Unfortunately, such a comprehensive framework that can respond to all the above questions does not exist in the maritime sector and any industry. Therefore, this study will address the following research questions raised from pending HRA issues in the maritime sector.

1) Is there a way to effectively identify human roles and better predict possible human errors in procedures or systems that change rapidly and become more complex?
2) Are factors that negatively or positively affect human performance suitably extracted, and are uncertainties and weights well addressed in measuring the level of extracted factors?
3) Is there an appropriate way to consider the interdependence of errors and interactions between components in quantifying human errors?
4) Is there a model integrating the predicted individual error probabilities into the risk assessment framework to obtain entire system reliability?
5) Is the HRA method used for each research project found to be appropriate for the analysis's aim and objectives?

3.3 Research aim and objectives

Developing a reliability assessment framework including human factors has a huge potential to innovate and improve the way safety is managed in maritime operations and systems. Therefore, the main aim of this study is to develop a more reliable and comprehensive human reliability assessment framework that can cover system reliability and be applied to various maritime operations and systems. Secondly, the framework for evaluating human reliability should be flexible enough to be optimised for various cases, depending on the complexity and interaction of operations, data availability, human and material resources, and the significance of failure.

Thus, by achieving the research aim, it is expected that this research enables the evaluation of the risk of systems and operations, including human factors, and can contribute to improving the safety of the design, process, and facilities at sea, such as system improvement and procedures improvement. In addition, an optimised HRA method provides a more practical and feasible integrated human risk analysis solution to the maritime industry.

To achieve the aforementioned overall aim, this research study will adopt the following specific objectives:

- To critically review the literature relevant to the current maritime human error prediction, human reliability, and system safety to identify the shortcomings of the current research and available methods.
- To derive or develop customised performance shaping factors suitable for each context arising from characteristics of the Maritime operations or systems.
- To modify and develop advanced human error quantification techniques
- Enhance human identification technique for complex system
- Enhance human error representation modelling to assess system reliability and integrate errors into PRA
- Establish criteria and a mechanism for determining the most appropriate HRA approaches for a particular project.
- Develop the following frameworks to provide optimised assessment ways for various analysis purposes and characteristics of the analysis target.
    1) Develop an instant human error calculation model that responds to immediate and straightforward analysis needs.
    2) Develop a human error calculation framework for extended human activities
    3) Develop a human reliability assessment framework to integrate human error into a probabilistic risk assessment framework that can cover the system reliability assessment
    4) Develop a human reliability assessment framework in a complex system by enhancing the human error identification approach

The final output of this research will be an integrated framework that will allow assessing human errors and integrating human errors into the risk picture to present the entire system's reliability. The study also provides individually optimised analysis methods for human error identification, quantification, and representation phase. Individual techniques can be flexibly combined and used for analysis purposes. Furthermore, evaluation criteria and framework will be provided to determine the optimal HRA method per project.

3.4 Chapter summary

This chapter has introduced the study's research questions, aim, and objectives. Chapter 4 will outline the approach adopted for this research.

4 Approach adopted

4.1 Chapter overview

This chapter presents the approach adopted to fulfil the aim and objectives of this research study. This study consists of four case studies about the maritime operation and system analysis presented in chapters 6, 7, 8 and 9, and four evaluations in chapter 5 to determine the optimal human reliability analysis method for each case study. Case studies from the maritime industry begin with identifying human error modes and quantifying human error, developing a system reliability model that incorporates human reliability, and assessing system reliability in complex systems. As a result, in the following order, the study methodology extends from the simple approach for determining human error modes to the comprehensive model for assessing system reliability.

1) Method and criteria to select the appropriate HRA method (Chapter 5)

2) Method to determine the human error mode (Chapter 6)

3) Human error quantification method for the human-oriented task (Chapter 7)

4) System reliability model that incorporates human error (Chapter 8)

5) Framework for the system reliability assessment in a complex system (Chapter 9)

4.2 Mind map of approach

The first phase in the HRA process is to define the scope and type of analysis and the responsibilities and human actions that will be assessed. Both qualitative and quantitative analysis is possible. The tasks being examined may be routine or emergency in nature. Two considerations must be taken into account while determining the scope of the research:

1) The analysis's objective of PRA may be to assist with an accident investigation, evaluate anomalies or issue reports, enhance processes or operations, or explore design trades.

2) The system's vulnerability to human error depends on its complexity, the degree to which the human interacts with it, and the degree to which the human-system interface is integrated.

As illustrated in ***Figure 4. 1***, a mind map was created to develop a straightforward approach and propose research activities to accomplish the study objectives for system reliability assessment, including human error in the maritime environment. The mind map depicts the work stages required for system reliability assessment, their particular processes, and the methodologies required to complete each step.

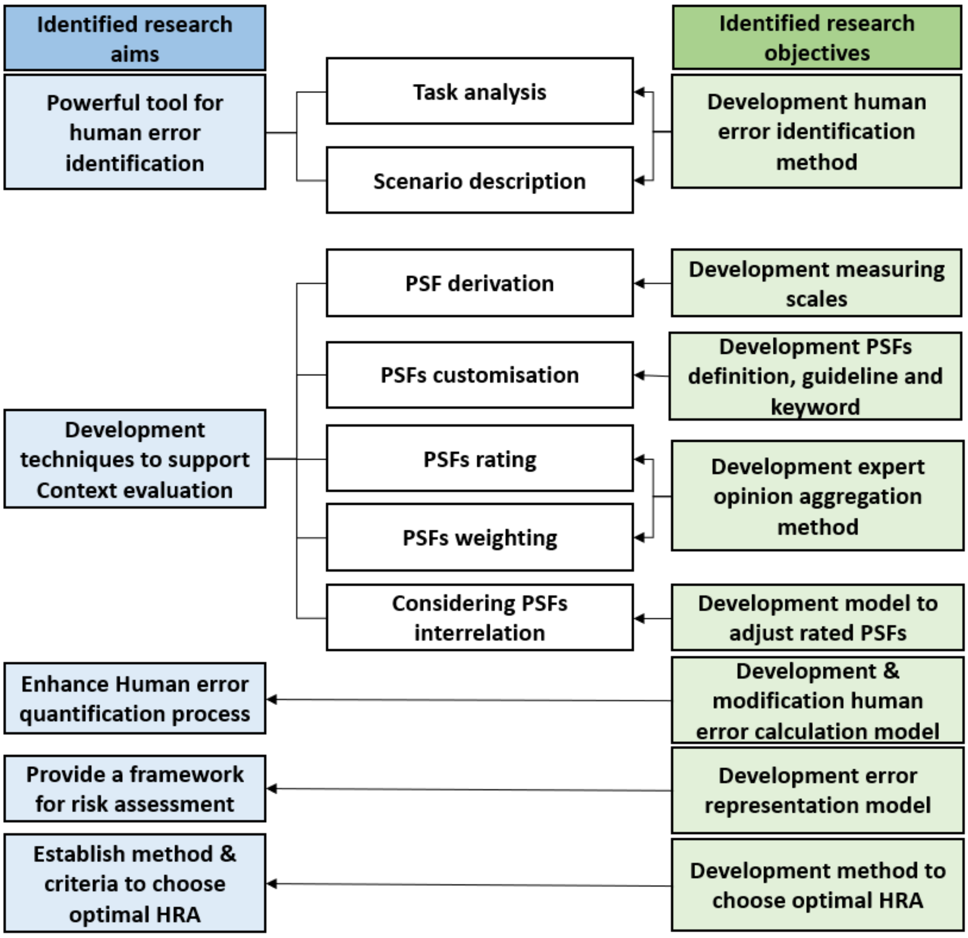

Figure 4. 1 Mind map of the approach adopted

4.2.1 Human error identification

It is essential to define and analyse the role of humans and the conditions surrounding them to identify possible human errors. Task analysis refers to methods of formally describing and analysing human-system interaction (Kirwan, 1994). Task analysis is conducted to define the steps which address the designated duties that the crew should complete successfully to achieve the main goal of the procedures with a hierarchical task analysis from the selected scenario. The purpose of task analysis in this research can be defined by simply subdividing the functions into tasks, tasks into subtasks, and subtasks into human actions. A task analysis describes the steps performed as part of the activity, providing a method of systematically organising the information collected about the task (Bye et al., 2017).

In this study, two different task analysis methods are utilised. First, a Hierarchical Task Analysis (HTA) is performed to define the task on the procedure's primary goal, along with subtasks to address the specified duties that the operator should complete. The HTA provides a graphical overview of the tasks involved in the analysis scenario. However, hierarchical task analysis is not sufficient to provide appropriate information in the context associated with the tasks. Therefore, Tabular Task Analysis (TTA) is also used as required to provide more details to expert judgment and better organise data.

However, human activity should not be treated as an independent element in complex and interactive dynamic systems or operations. Even when performing the same task, the action results can vary depending on the system situation and other variables such as the working environment and time. This is because the phase of the system changes dynamically as humans interact with machines and software. Thus, predicting possible human errors in these complex situations has limitations in traditional decomposition analysis. Therefore, this study proposes a human error identification method based on STPA, and its details are covered in Chapter 9.

4.2.2 Context evaluation and adjustment

In this study, context evaluation is defined as the process of identifying which factors affect human performance and measuring the degree of their influence. In addition, adjustment refers to the process of adjusting the size of the impact by the interrelationship of each influencing factor measured. Factors influencing human performance used in this study are presented in ***Table 4. 1***. The context evaluation process is described below to identify the methods required to achieve research goals for each phase.

First, the selection of PSFs that affect human performance and their assessment criteria should change depending on context. Therefore, the PSFs definition, provided by Whaley et al. (2011), was refined by maritime experts with customised guidance to establish and rate characteristics in specific operations. The provided description of PSFs should be as straightforward as possible for the experts to determine the appropriate PSFs rating for the task being analysed. Therefore, as shown in ***Table 4. 1***, the definition and selection of PSFs were chosen differently according to the desired operation and method. Once PSFs have been determined, the evaluation of PSFs needs to be conducted. This process is the most crucial step in quantification. As the result of the analysis relies on expert opinions, the deviation of expert selection, subjectivity, and bias should be minimised in this process. Mechanical selection should also be prevented, and the process should reflect an accurate expert evaluation. In addition, the measured value should be adjusted considering each PSF's relative importance and interdependence. To this end, the Fuzzy opinion aggregation method and model based on the Fuzzy theory and BN are suggested in this study.

Table 4. 1 PSFs labels per different HRA methods used in this study

Selected HRA		CREAM	SPAR-H	SLIM
Applied case studies		Engine room fire drill & Emergency steering	Rescue boat drill for Man overboard	LNG bunkering
PSF Categories	Mental	N/A	Threat Stress	N/A
	Organisation	Adequacy of organisation	Ship safety management system (SMS) and supports	Organisational factors
	Work environment	Working condition	The working condition	Working condition
	Man-machine interface	Adequacy of MMI and operational condition	Human-machine interface	Interface(Input device) Interface(Output device)
	Procedures	Availability of procedures/plan	Procedures	Procedure
	Work complexity	Number of simultaneous goals	The complexity of the task	Complexity
	Time	Available time	Time pressure	Time
	Environmental condition	Time of day	Environmental condition	Environmental condition
	Training & experience	Adequacy of training and experience	Level of experience or training skill	Experience & training
	Cooperation	Crew collaboration quality	N/A	N/A

4.2.3 Human error quantification

A total of four human error assessment techniques are presented in this study for quantifying human errors. Various human error quantification frameworks have been developed to provide the most efficient and optimised methods based on the characteristics of the chosen analysis and the objective rather than the use of a single method that fits all. Depending on the HRA selected, individual quantification methods are used in the human quantification process. Still, the basic concept appears in ***Figure 4. 2*** as human error quantification using a Fuzzy inference system. The values of PSF measured by experts or objective data are adjusted by the Fuzzy inference system and quantified into human errors. As part of the framework, BN-CREAM methods for immediate error calculations are addressed in Chapter 6, CREAM methods for human-oriented tasks are introduced in Chapter 7, SPAR-H methods are presented in Chapter 8, and SLIM methods are in Chapter 9.

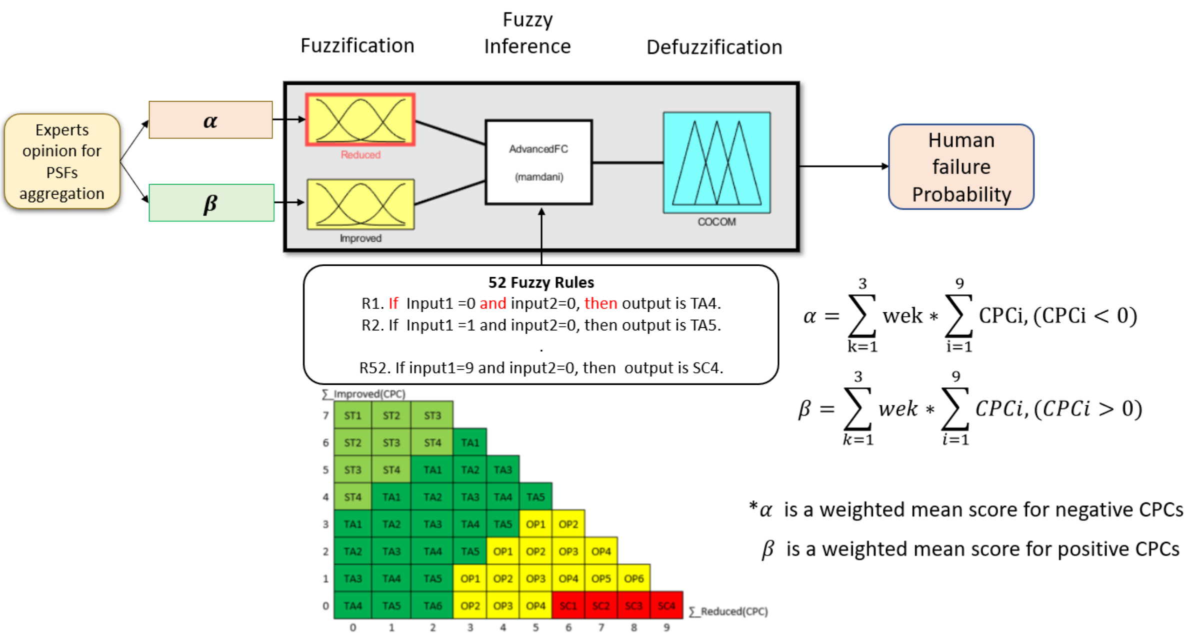

Figure 4. 2 Concept of the CREAM based Fuzzy inference system

4.2.4 Human error representation (The Integrated risk analysis methodology)

The objective of modelling human errors is to determine the total reliability levels inherent in the system, as shown in ***Figure 4. 3***, by adding up the risk probabilities of all failures and reliability combinations for hardware, software and humans. This is achieved by expressing it with human error in a logical tree known as the failure and event tree. However, this research proposes a new approach using the Reliability Block Diagram (RBD), assuming that each task and sub-work are system components for this HRA modelling. In particular, a new approach to modelling human reliability evaluation in individual human error probabilities based on reliability block diagrams can be applied to various systems. In addition, through this approach, the relationship between complex tasks that cannot be achieved only by hierarchical task analysis is effectively expressed in a simplified manner. Details will be explained in conjunction with case studies illustrated in Chapters 8 and 9 on rescue boat drill and LNG bunkering processes, respectively.

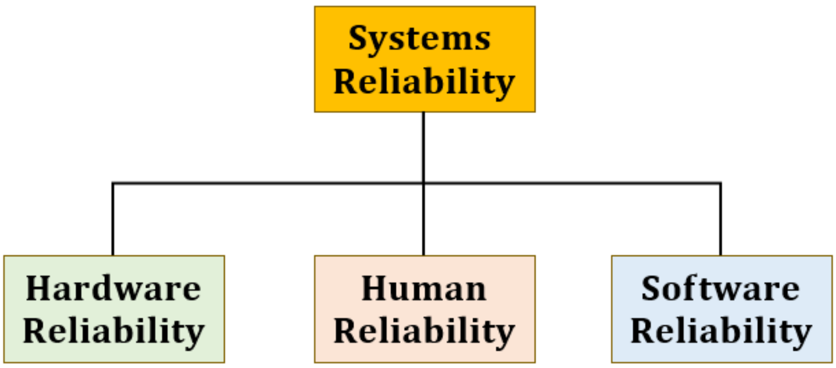

Figure 4. 3 Abstraction of system reliability

4.2.5 Determine the appropriate HRA method for a given project

Human reliability is the most unpredictable and accounts for a significant part of system reliability research. For this human reliability assessment, there are various HRA methods from which to choose, but selecting the one that best suits a given research situation is a crucial first step for individual case studies. As a result, a method for establishing criteria for evaluating individual HRA and deriving the results must be developed. Selecting the most appropriate HRA method for each case study will be detailed in Chapter 5.

4.3 Overall framework for research

This section briefly summarises and introduces the work phase and the corresponding methodology to perform each step. ***Figure 4. 4*** depicts an overall framework that clarifies how this research study's processes and outcomes are connected and interrelated. Further details of each methodology will be covered in the specific case studies from Chapters 5 to 9.

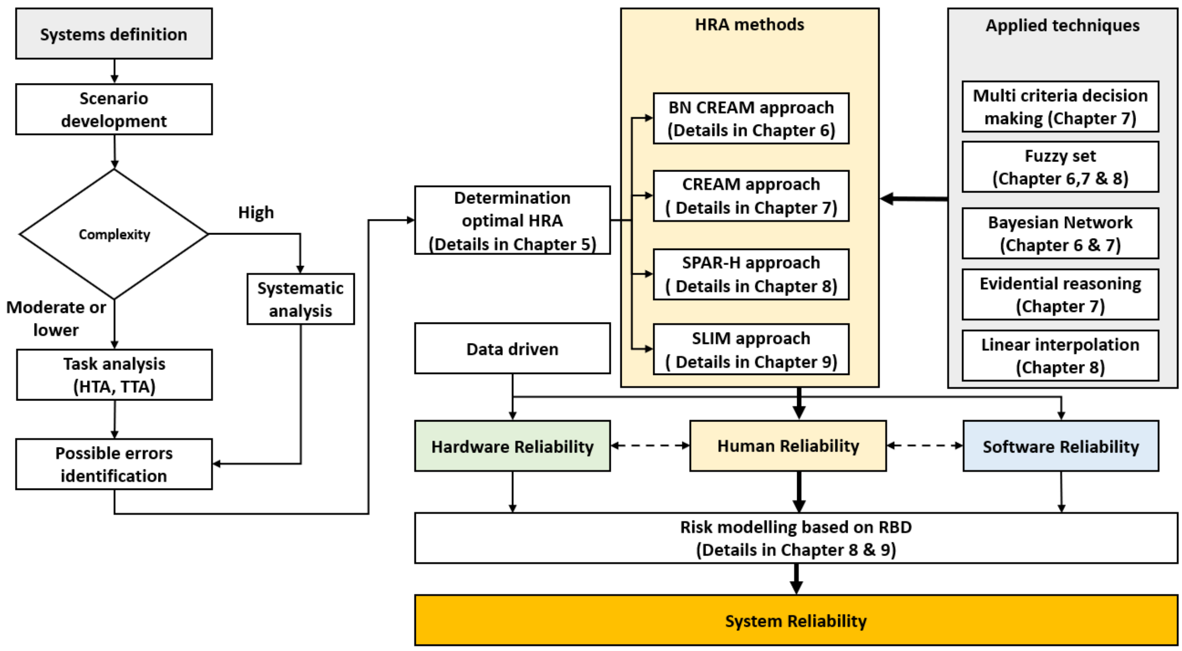

Figure 4. 4 Overall frameworks of the study

4.4 Chapter summary

This chapter presents the approaches adopted for this study. Each methodology in Chapters 5 to 9 will be presented explicitly with the case studies.

5 Determining the optimal HRA method

5.1 Chapter overview

Although numerous HRA techniques are available, no study has attempted to find the ideal HRA techniques for maritime systems. This chapter aimed to select an appropriate human reliability analysis method from a pool of available methods suitable for the research circumstances. Human reliability analysis methods vary in terms of their traits and capabilities and the extent of their application. As a result, it is vital to develop criteria for evaluating HRA approaches to determine the best appropriate method for the research goal. For evaluating HRA approaches, four criteria and fourteen sub-criteria for selecting the ideal method were determined through a review of the literature and expert opinion. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method by Yoon and Hwang (1995) was used to prioritise the HRA methods from a pair of the six HRA methods. Four assessments were conducted on the case studies that comprise this thesis.

5.2 Criteria identified for determination of HRA methods

The criteria and sub-criteria of the research were determined based on previous studies, including Abbaspour et al. (2020), Bell and Holroyd (2009) and Mosleh et al. (2006). The criteria consist of four aspects: qualitative analysis capacities, quantitative analysis capacities, resource requirements, and task complexity. The selected four criteria and their 14 sub-criteria are presented to compare HRA methods in ***Table 5. 1***. The Likert scales for each sub-criteria were phrased to avoid confusion and increase the effectiveness in assessing criteria. Likert scales are generally used to measure changes in attitude, knowledge, perception, value, and behaviour because the Likert scale includes a series of statements that respondents can choose to evaluate responses to evaluation questions (Vogt and Johnson, 2011). The cost criteria, requiring knowledge and consuming time, are among the 14 sub-criteria, whereas the other criteria are benefit criteria. Individual Likert scales are developed and scored on a scale of 1 to 5, such that each sub-criteria becomes a benefit criterion. In other words, from the standpoint of the study's benefit, the scale with the most significant profit is labelled 5, while the scale with the most loss is labelled 1. For instance, the lower the required level of knowledge, as determined by the Cost Criterion, the more beneficial it is to the whole study plan; the lowest level is 5. However, the higher the number, the greater the benefit to the total study. Scales all have the same direction, regardless of the type of criterion utilised. This is to minimise the number of unnecessary calculations.

Table 5. 1 Criteria to evaluate HRA methods

Sub-criteria	Description
	C1 Qualitative analysis capacities
SC1 Guidance on task decomposition	When using the HRA approach, task decomposition is an essential step. It involves breaking down human activities of interest into sub-tasks congruent with the method's "analytical unit" or "basic tasks." Guidelines for these task breakdown processes are necessary and offered depending on the HRA, but no alternative is provided. As a result, this criterion aims to determine how effective the selected HRA method is at performing task decomposition. The evaluation scale is based on the following factors: the number of different types of human behaviour provided, the amount of quantitative and qualitative guidance supplied, and the degree of diversity of human behaviour provided.
SC2 Performance Shaping Factors (PSFs) list	Human error is associated directly or indirectly with several elements referred to as performance shaping factors (PSFs). These PSFs are characteristics of human behaviour and situations that can influence human performance. They are frequently used to predict human error probability (HEPs) and identify contributors to human performance. Performance shaping factors can be used to assess situations from various viewpoints. If additional PSF can be available, it benefits HRA analysis since the presence and effectiveness of additional PSF may be determined through observations, interviews, and existing records. This criterion was created to assess applicability regarding the number of PSFs and their accompanying rules for each HRA approach. The rating scale is based on the number of distinct types of PSFs supplied and the quantitative and qualitative features of the provided PSF guidelines.
SC3 PSFs coverage	The first HRA technique established a quantitative risk analysis that considered human behaviour and errors. However, these methodologies ignore cognitive features, commission errors, contexts, and organisational issues when analysing PSF. In contrast, the second generation HRA focuses on human cognitive aspects, and the first generation HRA focuses exclusively on behavioural components of human reliability. This criterion was used to determine whether the given PSF could handle the additional cognitive, ergonomic, and organisational requirements of more sophisticated modern systems.

SC4 PSF flexibility	Most human reliability analysis approaches use predetermined PSF sets for either predictive or retrospective analysis. The analysers can specify the PSF set following the analysis task. However, PSFs that influence maritime operations are distinct from those that affect onshore operations. Because of this, the HRA technique should be capable of adapting to the unique PSFs of the maritime industry through customisation or extension of the PSF set to represent the working environment and mission characteristics specific to the maritime sector.
SC5 Applicability to maritime	The term "applicability" refers to the degree to which the chosen HRA approach applies to various activities and systems during critical operations and emergency response scenarios, including ordinary duties performed in the maritime sector. This is a vital criterion given the study's primary focus on the maritime industry.
C2 Quantitative analysis capabilities	
SC6 Nominal HEP for specific error mode	When human error is quantified, the HRA approach determines the nominal human error probability assigned to a given type of human behaviour. More human behaviour may be materialised by including a range of nominal human error probability, resulting in more realistic error values. This criterion seeks to ascertain the appropriate degree of the nominal error probability.
SC7 Sensitivity of rated PSFs	PSF is classified as direct and indirect PSF. As a result, the sensitivity of individual PSFs to impact human performance should be established correctly based on their features. This criterion is used to determine the suitability of PSF sensitivity. While increased sensitivity does not necessarily imply superiority, experts should carefully analyse each study project.
SC8 Diversity of analysis approach	The diversity of analysis approaches allows for additional flexibility in delivering several methods depending on whether the analysis is performed manually or with software assistance.
C3 Resource requirements	
SC9 Required Knowledge Level for HRA (usability)	The level of knowledge required for HRA denotes the level of expertise that the analyst should possess when performing the specified HRA procedure. As a scale of measurement, it can be determined whether it is suitable for the experienced workers in a specific industry with brief education/introduction, whether a general human reliability analyst can perform it, or whether it requires the expertise of a specialised human reliability expert.
SC10 Required level of effort / consuming time	The required level of effort/consuming time denotes the expected time and effort required to complete the analysis. This is also the primary factor in real-world research situations.

SC11 Level of familiarisation for selected HRA	The level of familiarity with the specified HRA technique indicates the level of knowledge of the presently secured analyst with the selected HRA approach. This criterion quantifies the degree to which research team members are prepared to conduct HRA analysis. This criterion does not reflect the analyst's overall HRA knowledge but the specific HRA skill and expertise that should be evaluated.
C4 Complexity of the task	
SC12 Processable System complexity	The complexity of a system that can be processed is defined as the degree of complexity that the chosen HRA can manage. The degree of interaction between humans, machines, and software quantifies system complexity.
SC13 Processable Communication complexity	Processable communication complexity refers to the level of complexity that the HRA chosen can manage. Communication complexity is quantified by the number of persons involved in the operation or system and the number of independent parties.
SC14 Convertibility to other methods	While most HRA techniques focus on measuring human errors, they frequently only give limited or no identification of human errors or risk assessment frames. As a result, integration with distinct methodologies is required. Thus, convertibility serves as a criterion for determining how easily a selected HRA may be integrated with other models or methods.

5.3 Methodology

To determine optimal HRA method under the given circumstance, the framework based on the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) was developed to assess HRA methods compatible with maritime systems in ***Figure 5. 1***. The TOPSIS is a multi-criteria decision analysis method and constitutes two a-phase solutions. One solution means an Ideal Solution, a virtual best solution made with only the available information in the decision matrix, and the other is, on the contrary, the worst solution found with only the most Negative Ideal Solution, that is, the worst solution found with only the available information. After creating these two ideal solutions, the best alternative is the closest to the ideal solution and the farthest from the negative ideal solution. Here, this study adopts the city-block distance method as the distance measure. Because, unlike the Euclidean distance method, the city-block distance method ensures that the closest alternative to the ideal solution is the farthest from the negative ideal solution. The steps of TOPSIS in this study based on multi-attribute decision-making presented by Yoon and Hwang (1995) to determine the optimal HRA method are as follows.

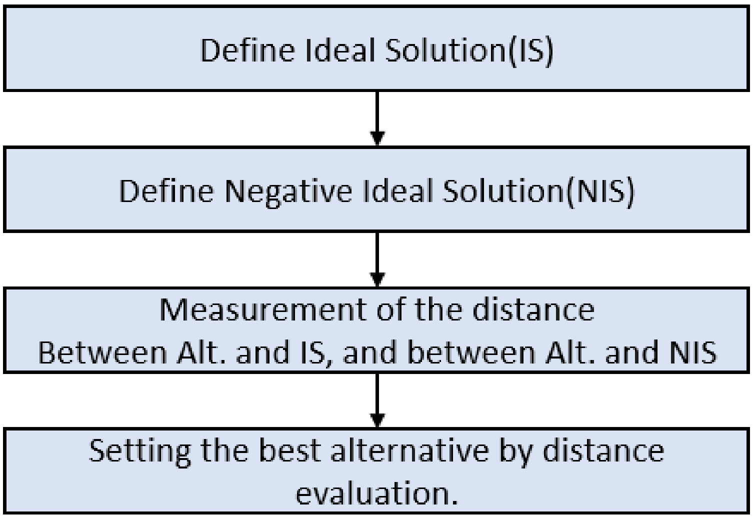

Figure 5. 1 Overall flowchart of TOPSIS (Adopted from Yoon and Hwang (1995))

Step 1: Create an evaluation matrix

Create an evaluation matrix consisting of m alternatives (A_m) and n criteria (C_n), with the intersection of each alternative and criteria given as $x_{ij,}$ to have a matrix $(x_{ij})_{m \times n}$ as bellows.

$$
\begin{array}{c|cccc|}
 & \mathbf{C_1} & \mathbf{C_2} & & \mathbf{C_n} \\
\mathbf{A_1} & x_{11} & x_{12} & & x_{1n} \\
\mathbf{A_2} & x_{21} & x_{22} & & . \\
\mathbf{A_3} & x_{31} & . & \ldots\ldots & . \\
. & . & . & & . \\
. & . & . & & . \\
. & . & . & & . \\
\mathbf{A_m} & x_{m1} & . & \ldots\ldots & x_{mn}
\end{array}
$$

Step 2: Normalisation of the Evaluation matrix

The matrix $(x_{ij})_{m \times n}$ is then normalised to form the matrix $(r_{ij})_{m \times n}$ using the Vector normalisation method as follows.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}{}^{2}}} \qquad \text{Where i= 1, 2,....,m, j=1, 2,...,n}$$

Step 3: Calculate weighted normalised matrix

The weighted normalised matrix refers to a matrix in which comparable dimensionless data and weights are considered. The weighted normalised decision matrix can be calculated as follows.

$$v_{ij} = r_{ij} \times w_j$$

Where $w_j = \frac{W_j}{\sum_{k=1}^{n} W_k}$ where j=1, 2,…n, so that $\sum_{i=1}^{n} W_i = 1$,

and W_j is the original weight given to the indicator v_j, j=1, 2...,n

Step 4: Determine the Ideal Solution (A^b) and the Negative Ideal solution (A^w)

Each criterion of the weighted normalisation matrix is classified into benefit criterion(J^+) and cost criterion(J^-). The benefit criteria show a positive value in big numbers, whereas the cost-benefit criteria indicate a positive value in small numbers. Accordingly, the ideal and negative ideal solutions may be made as follows.

$$A^b = \{min(v_{ij} | i = 1,2,\dots,m) | \, j \in J_-\}, \{max(v_{ij} | i = 1,2,\dots,m) | \, j \in J_+\}$$

$$\equiv \{v_{bj} | j = 1,2,\dots,n\}$$

$$A^w = \{max(v_{ij} | i = 1,2,\dots,m) | \, j \in J_-\}, \{min(v_{ij} | i = 1,2,\dots,m) | \, j \in J_+\}$$

$$\equiv \{v_{wj} | j = 1,2,\dots,n\}$$

Step 5. Distance calculation

Calculate the distance (S_i^-) between the target alternative (A_i) and the worst condition (A^w) and the distance (S_i^+) between the alternative (A_i) and the best condition (A^b) as follows.

$$S_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_{wj})^2}\,, \text{ and } S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_{bj})^2}\,, \text{ i = 1, 2, ..., m}$$

Step 6: Calculate the similarity to the worst condition and rank alternatives

According to the City-block distance method, the Similarity (C_i) is obtained as follows, and the alternative with the largest Similarity (C_i) means the optimal alternative.

$$C_i = \frac{S_i^-}{(s_i^- + s_i^+)}$$

5.4 Assessing and ranking the alternatives per case study

To assess and compare HRA approaches, the TOPSIS questionnaire, which was developed based on the methods and sub-criteria evaluated, was used to collect expert judgments from participants. At this point, an attempt was made to acquire the opinions of specialists on each of the HRA methods under consideration. To ensure this phase was successful, HRA experts who were thoroughly familiar with and had a strong command of at least one of the previously described HRA approaches were invited to participate. The subject matter experts provided adequate information about the study's goal and approach. They were asked to respond to a questionnaire on the subject. The HEART, THERP, SLIM, SPAR-H, CREAM, and BN-CREAM methodologies were all selected to conduct a comparative analysis. ***Table 5. 2*** summarises the fundamental aspects of the HRA approaches that have been set. More information on the selected HRAs can be found in Chapter 2 of this thesis. Each research project included in this thesis was evaluated to determine the most appropriate HRA method. However, this result does not mean that a specific HRA method has an absolute advantage or inferiority over other methods. It is a method of finding the most reasonable and efficient alternative based on resources such as the scope, available period, and human resources of the study under given conditions. Because of this, the same HRA method may give different results depending on when the evaluation is done and what the context is. The following case studies were evaluated separately to determine which HRA analysis method is most appropriate for the four research tasks chosen for this thesis.

Table 5. 2 Characteristics of the selected HRA methods

HRA	Task decomposition	Performance Shaping Factors (PSFs) list	Nominal HEP for specific error mode	Diversity of Analysis Approach
BN-CREAM	15 cognitive activities	9 CPCs	Not required	Manual and Software aided
CREAM	15 cognitive activities	9 CPCs	13 Credible failure types	Manual only
SPAR-H	Two types of tasks; Diagnosis, action	8 PSFs for quantification, many for root causes	Diagnosis, Action	Manual only
SLIM	None specified	8 PSFs	None specified (End anchoring method)	Manual only

THERP	27 types of activities	5 PSFs and dependency relations	14 HEP Error modes and HRA event tree	Manual only
HEART	8 Generic tasks	38 Error Producing Conditions (EPCs)	8 Nominal HEP for GTT	Manual only

5.4.1 Determining optimal HRA for case study 1: emergency steering case (evaluation 1)

This section describes the procedure for determining the most appropriate approach for doing human reliability analysis during emergency steering. First, section 5.4.1.1 summarises the project's fundamentals; specifics will be presented in Chapter 6. Then, expert analysis was undertaken using the provided data to choose the HRA approach that is most appropriate for the study aim. Five experts participated, and the input values represented the workshop participants' consensus viewpoints.

5.4.1.1 Project description of emergency steering project

This research aims to discover the error mode by analysing the components that affect emergency steering operation when the ship's primary propulsion system fails and then estimate the total probability of human error. Individual task analysis is not required, and the goal is to achieve the research purpose in a relatively short time. Therefore, integration with risk analysis is not necessary, and emergency steering necessitates communication between the steering room and the control room, as well as consideration of human-machine interaction, but not software. However, due to the limited labour and time resources available, analysis by software is essential.

5.4.1.2 Application of TOPSIS to determine optimal HRA for emergency steering project

Table 5. 3 summarises the evaluation results for the six HRA methods with scores ranging from 1 to 5 for this project's 14 evaluation criteria. According to the methods described in Section 5.3, normalised assessment metrics were derived as shown in ***Table 5. 4***. ***Table 5. 5*** summarises the relative importance of each criterion.

Table 5. 3 Evaluation matrix for emergency steering

Alternatives	**Criteria**													
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
BN-CREAM (A_1)	5	4	4	3	4	4	3	4	5	5	5	2	2	2
CREAM (A_2)	5	4	4	3	4	4	3	3	3	3	3	3	3	3
SPAR-H (A_3)	3	3	2	4	3	2	4	3	2	3	3	3	3	3
SLIM (A_4)	3	3	3	4	4	2	3	3	3	3	3	3	3	4
THERP (A_5)	4	2	4	2	2	3	3	3	1	1	1	4	2	2
HEART (A_6)	4	5	3	2	2	3	3	3	1	2	2	4	2	2

Table 5. 4 Normalised evaluation matrix for emergency steering

Alternatives	**Criteria**						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.5000	0.4500	0.4781	0.3939	0.4961	0.5252	0.3841
CREAM	0.5000	0.4500	0.4781	0.3939	0.4961	0.5252	0.3841
SPAR-H	0.3000	0.3375	0.2390	0.5252	0.3721	0.2626	0.5121
SLIM	0.3000	0.3375	0.3586	0.5252	0.4961	0.2626	0.3841
THERP	0.4000	0.2250	0.4781	0.2626	0.2481	0.3939	0.3841
HEART	0.4000	0.5625	0.3586	0.2626	0.2481	0.3939	0.3841
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.5121	0.7143	0.6623	0.6623	0.2520	0.3203	0.2949
CREAM	0.3841	0.4286	0.3974	0.3974	0.3780	0.4804	0.4423
SPAR-H	0.3841	0.2857	0.3974	0.3974	0.3780	0.4804	0.4423
SLIM	0.3841	0.4286	0.3974	0.3974	0.3780	0.4804	0.5898
THERP	0.3841	0.1429	0.1325	0.1325	0.5040	0.3203	0.2949
HEART	0.3841	0.1429	0.2649	0.2649	0.5040	0.3203	0.2949

Table 5. 5 Weights and normalised weights for criteria for emergency steering

Criteria	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**
W_j	3	4	3	2	5	5	3
w_j	0.07	0.09	0.07	0.05	0.11	0.11	0.07
Criteria	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**
W_j	2	4	4	4	2	2	1
w_j	0.05	0.09	0.09	0.09	0.05	0.05	0.02

Table 5. 6 Weighted normalised evaluation matrix for emergency steering

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.0341	0.0409	0.0326	0.0179	0.0564	0.0597	0.0262
CREAM	0.0341	0.0409	0.0326	0.0179	0.0564	0.0597	0.0262
SPAR-H	0.0205	0.0307	0.0163	0.0239	0.0423	0.0298	0.0349
SLIM	0.0205	0.0307	0.0244	0.0239	0.0564	0.0298	0.0262
THERP	0.0273	0.0205	0.0326	0.0119	0.0282	0.0448	0.0262
HEART	0.0273	0.0511	0.0244	0.0119	0.0282	0.0448	0.0262
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.0233	0.0649	0.0602	0.0602	0.0115	0.0146	0.0067
CREAM	0.0175	0.0390	0.0361	0.0361	0.0172	0.0218	0.0101
SPAR-H	0.0175	0.0260	0.0361	0.0361	0.0172	0.0218	0.0101
SLIM	0.0175	0.0390	0.0361	0.0361	0.0172	0.0218	0.0134
THERP	0.0175	0.0130	0.0120	0.0120	0.0229	0.0146	0.0067
HEART	0.0175	0.0130	0.0241	0.0241	0.0229	0.0146	0.0067

Based on ***Table 5. 6***, the best and worst alternatives are obtained as follows.

The worst alternative $\mathbf{A^w}$ = (0.02045, 0.02045, 0.01630, 0.01194, 0.02819, 0.02984, 0.02619, 0.01746, 0.01299, 0.01204, 0.01204, 0.01145, 0.01456, 0.00670) and

The best alternative $\mathbf{A^b}$ = (0.03409, 0.05114, 0.03260, 0.02387, 0.05638, 0.05968, 0.03492, 0.02328, 0.65000,0.06021, 0.06021, 0.02291, 0.02184, 0.01340)

The distance (S_i^-) between the target alternative (A_i) and the worst condition (A^w) and the distance (S_i^+) between the alternative (A_i) and the best condition (A^b) were determined in ***Table 5. 7***.

Table 5. 7 The distances of S_i^- and S_i^+ and the rank of alternatives for emergency steering

Alternatives	S_i^-	S_i^+	Similarity	Rank
$\boldsymbol{A_1}$	**0.10**	**0.59**	**0.15**	**1**
A_2	0.07	0.61	0.10	2
A_3	0.04	0.63	0.07	4
A_4	0.06	0.61	0.08	3
A_5	0.03	0.64	0.04	6
A_6	0.04	0.64	0.06	5

In conclusion, the Bayesian network (BN) CREAM approach was determined to be the most effective way for the emergency steering research project.

5.4.2 Determining optimal HRA for case study 2: engine room fire drill case (evaluation 2)

This section describes the procedure for determining the most appropriate approach for doing human reliability analysis during engine room fire drill procedures. Section 5.4.2.1 summarises the project's fundamentals; specifics will be presented in Chapter 7. It was examined in the same manner as section 5.4.1.

5.4.2.1 Project description of engine room fire drill project

This study aims to evaluate human error per each task in emergency response drills to identify the most vulnerable human activities and prioritise human duties during ship engine room fires. Individual tasks must be determined via task analysis, as well as the connection between humans and machines. Numerous sites must be analysed, including steering rooms, fire scenes, and fire control rooms. The described task is centred on the human-oriented task and demands execution and cognitive activity. The project will take around six months and require no risk assessment analysis.

5.4.2.2 Application of TOPSIS to determine optimal HRA for engine room fire drill project

Table 5. 8 summarises the evaluation results for the six HRA methods with scores ranging from 1 to 5 for this project's 14 evaluation criteria. According to the methods described in Section 5.3, a normalised evaluation matrix was derived as shown in ***Table 5. 9***.

***Table 5.* 10** summarises the relative importance of each criterion for engine room fire drill procedures.

Table 5. 8 Evaluation matrix for engine room fire

Alternatives	Criteria													
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
BN-CREAM (A_1)	5	4	4	3	4	4	3	4	5	5	5	2	2	2
CREAM (A_2)	5	4	4	3	4	4	3	3	3	3	4	3	3	3
SPAR-H (A_3)	3	3	2	4	3	2	4	3	2	3	3	3	3	3
SLIM (A_4)	3	3	3	4	4	2	3	3	3	3	3	3	3	4
THERP (A_5)	4	2	4	2	2	3	3	3	1	1	1	4	2	2
HEART (A_6)	4	5	3	2	2	3	3	3	1	2	2	4	2	2

Table 5. 9 Normalised evaluation matrix for engine room fire

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.50000	0.45004	0.47809	0.39392	0.49614	0.52523	0.38411
CREAM	0.50000	0.45004	0.47809	0.39392	0.49614	0.52523	0.38411
SPAR-H	0.30000	0.33753	0.23905	0.52523	0.37210	0.26261	0.51215
SLIM	0.30000	0.33753	0.35857	0.52523	0.49614	0.26261	0.38411
THERP	0.40000	0.22502	0.47809	0.26261	0.24807	0.39392	0.38411
HEART	0.40000	0.56254	0.35857	0.26261	0.24807	0.39392	0.38411
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.51215	0.71429	0.66227	0.62500	0.25198	0.32026	0.29488
CREAM	0.38411	0.42857	0.39736	0.50000	0.37796	0.48038	0.44233
SPAR-H	0.38411	0.28571	0.39736	0.37500	0.37796	0.48038	0.44233
SLIM	0.38411	0.42857	0.39736	0.37500	0.37796	0.48038	0.58977
THERP	0.38411	0.14286	0.13245	0.12500	0.50395	0.32026	0.29488
HEART	0.38411	0.14286	0.26491	0.25000	0.50395	0.32026	0.29488

Table 5. 10 Weights and normalised weights for criteria for emergency steering

Criteria	C1	C2	C3	C4	C5	C6	C7
W_j	4	4	3	2	5	5	3
w_j	0.08	0.08	0.06	0.04	0.10	0.10	0.06
Criteria	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**
W_j	2	3	3	4	4	4	2
w_j	0.04	0.06	0.06	0.08	0.08	0.08	0.04

Table 5. 11 Weighted normalised evaluation matrix for engine room fire

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.04167	0.03750	0.02988	0.01641	0.05168	0.05471	0.02401
CREAM	0.04167	0.03750	0.02988	0.01641	0.05168	0.05471	0.02401
SPAR-H	0.02500	0.02813	0.01494	0.02188	0.03876	0.02736	0.03201
SLIM	0.02500	0.02813	0.02241	0.02188	0.05168	0.02736	0.02401
THERP	0.03333	0.01875	0.02988	0.01094	0.02584	0.04103	0.02401
HEART	0.03333	0.04688	0.02241	0.01094	0.02584	0.04103	0.02401
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.02134	0.04464	0.04139	0.05208	0.02100	0.02669	0.01229
CREAM	0.01600	0.02679	0.02483	0.04167	0.03150	0.04003	0.01843
SPAR-H	0.01600	0.01786	0.02483	0.03125	0.03150	0.04003	0.01843
SLIM	0.01600	0.02679	0.02483	0.03125	0.03150	0.04003	0.02457
THERP	0.01600	0.00893	0.00828	0.01042	0.04200	0.02669	0.01229
HEART	0.01600	0.00893	0.01656	0.02083	0.04200	0.02669	0.01229

Based on **Table 5. 11**, the best and worst alternatives are obtained as follows.

The worst alternative A^w = (0.02500, 0.01875, 0.01494, 0.01094, 0.02584, 0.02736, 0.02401, 0.01600, 0.00893, 0.00828, 0.01042, 0.02100, 0.02669, 0.01229) and

The best alternative A^b = (0.04167, 0.04688, 0.02988, 0.02188, 0.05168, 0.05471, 0.03201, 0.02134, 0.04464,0.04139, 0.05208, 0.04200, 0.04003, 0.02457)

The distance (S_i^-) between the target alternative (A_i) and the worst condition (A^w) and the distance (S_i^+) between the alternative (A_i) and the best condition (A^b) were determined in ***Table 5. 12***.

Table 5. 12 The distances of S_i^- and S_i^+ and rank of alternatives for engine room fire

Alternatives	S_i^-	S_i^+	Similarity	Rank
A_1	0.08	0.05	0.60	2
A_2	0.06	0.04	0.63	1
A_3	0.04	0.06	0.40	4
A_4	0.05	0.05	0.48	3
A_5	0.03	0.08	0.27	6
A_6	0.04	0.07	0.38	5

In conclusion, the CREAM approach was determined to be the most effective way for the engine room fire drill research project.

5.4.3 Determining optimal HRA for case study 3: man overboard drill case (evaluation 3)

This section describes the procedure for determining the most appropriate approach for system reliability analysis, including human error, for rescue boat drills to man overboard on ships. Section 5.4.3.1 summarises the project's fundamentals; specifics will be presented in Chapter 8. It was examined in the same manner as section 5.4.1.

5.4.3.1 Project description of man overboard drill

This study aims to quantify human errors associated with each task in emergency rescue boat drills in the case of man overboard on ships and then integrate them into a risk model to determine system reliability. Individual tasks must be defined through task analysis and human-machine interaction. Numerous locations should be investigated, including the steering room, the inside of the rescue boat, and the rescue boat davit operating site. Not only should the specified task be human-oriented, but it should also account for human-machine interaction. This project will take approximately a year to complete and require additional risk assessment analysis, including machine failures. However, because a distinct risk analysis model based on a reliability block diagram will be employed, quantifying human error is critical for choosing the HRA method, but easy integration and compatibility with other approaches should also be addressed.

5.4.3.2 Application of TOPSIS to determine optimal HRA for man overboard drill project

Table 5. 13 summarises the evaluation results for the six HRA methods with scores ranging from 1 to 5 for this project's 14 evaluation criteria. According to the methods described in Section 5.3, a normalised evaluation matrix was derived in ***Table 5. 14***. ***Table 5. 15*** summarises the relative importance of each criterion for man overboard drill procedures.

Table 5. 13 Evaluation matrix for man overboard

Alternatives	Criteria													
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
BN-CREAM (A_1)	3	4	3	3	4	4	3	4	5	5	4	2	2	2
CREAM (A_2)	3	4	3	3	4	4	3	3	3	3	3	3	3	3
SPAR-H (A_3)	3	3	4	4	3	2	4	3	3	3	5	3	4	5
SLIM (A_4)	3	3	3	4	4	2	3	3	3	3	2	3	3	4
THERP (A_5)	4	2	4	2	2	3	3	3	1	1	1	4	2	2
HEART (A_6)	4	5	3	2	2	3	3	3	1	2	2	4	2	2

Table 5. 14 Normalised evaluation matrix for man overboard

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.36380	0.45004	0.36380	0.39392	0.49614	0.52523	0.38411
CREAM	0.36380	0.45004	0.36380	0.39392	0.49614	0.52523	0.38411
SPAR-H	0.36380	0.33753	0.48507	0.52523	0.37210	0.26261	0.51215
SLIM	0.36380	0.33753	0.36380	0.52523	0.49614	0.26261	0.38411
THERP	0.48507	0.22502	0.48507	0.26261	0.24807	0.39392	0.38411
HEART	0.48507	0.56254	0.36380	0.26261	0.24807	0.39392	0.38411
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.51215	0.68041	0.66227	0.52076	0.25198	0.29488	0.25400
CREAM	0.38411	0.40825	0.39736	0.39057	0.37796	0.44233	0.38100
SPAR-H	0.38411	0.40825	0.39736	0.65094	0.37796	0.58977	0.63500
SLIM	0.38411	0.40825	0.39736	0.26038	0.37796	0.44233	0.50800
THERP	0.38411	0.13608	0.13245	0.13019	0.50395	0.29488	0.25400
HEART	0.38411	0.13608	0.26491	0.26038	0.50395	0.29488	0.25400

Table 5. 15 Weights and normalised weights for criteria for man overboard

Criteria	C1	C2	C3	C4	C5	C6	C7
W_j	3	3	3	4	5	4	4
w_j	0.06	0.06	0.06	0.08	0.10	0.08	0.08
Criteria	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**
W_j	1	2	3	4	4	4	5
w_j	0.02	0.04	0.06	0.08	0.08	0.08	0.10

Table 5. 16 Weighted normalised evaluation matrix for man overboard

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.02227	0.02755	0.02227	0.03216	0.05063	0.04288	0.03136
CREAM	0.02227	0.02755	0.02227	0.03216	0.05063	0.04288	0.03136
SPAR-H	0.02227	0.02066	0.02970	0.04288	0.03797	0.02144	0.04181
SLIM	0.02227	0.02066	0.02227	0.04288	0.05063	0.02144	0.03136
THERP	0.02970	0.01378	0.02970	0.02144	0.02531	0.03216	0.03136
HEART	0.02970	0.03444	0.02227	0.02144	0.02531	0.03216	0.03136
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.01045	0.02777	0.04055	0.04251	0.02057	0.02407	0.02592
CREAM	0.00784	0.01666	0.02433	0.03188	0.03085	0.03611	0.03888
SPAR-H	0.00784	0.01666	0.02433	0.05314	0.03085	0.04814	0.06480
SLIM	0.00784	0.01666	0.02433	0.02126	0.03085	0.03611	0.05184
THERP	0.00784	0.00555	0.00811	0.01063	0.04114	0.02407	0.02592
HEART	0.00784	0.00555	0.01622	0.02126	0.04114	0.02407	0.02592

Based on ***Table 5. 16***, the best and worst alternatives are obtained as follows.

The worst alternative A^w = (0.02227, 0.01378, 0.02227, 0.02144, 0.02531, 0.02144, 0.03136, 0.00784, 0.00583, 0.00811, 0.01081, 0.02057, 0.02407, 0.02592) and

The best alternative A^b = (0.02970, 0.03444, 0.02970, 0.04288, 0.05063, 0.04288, 0.04181, 0.01045, 0.02777,0.04055, 0.05314, 0.04114, 0.04814, 0.06480)

The distance (S_i^-) between the target alternative (A_i) and the worst condition (A^w) and the distance (S_i^+) between the alternative (A_i) and the best condition (A^b) were determined in ***Table 5. 17***.

Table 5. 17 The distances of S_i^- and S_i^+ and rank of alternatives for man overboard

Alternatives	S_i^-	S_i^+	Similarity	Rank
A_1	0.06	0.05	0.53	2
A_2	0.05	0.05	0.53	3
A_3	0.07	0.04	0.66	1
A_4	0.05	0.05	0.49	4
A_5	0.03	0.08	0.23	6
A_6	0.03	0.07	0.32	5

In conclusion, the SPAR-H approach was determined to be the most effective way for the man overboard research project.

5.4.4 Determining optimal HRA for case study 4: ESD system during STS LNG bunkering case (evaluation 4)

This section describes the procedure for determining the most appropriate system reliability analysis, including human error, for the ESD system during the ship-to-ship LNG bunkering process. Section 5.4.4.1 summarises the project's fundamentals; specifics will be presented in Chapter 9. It was examined in the same manner as section 5.4.1.

5.4.4.1 Project description of ESD system during STS LNG bunkering

This study aims to analyse system reliability in complicated future designs, including humans, machines, and software. This section will identify the optimal HRA analysis technique that will be used to anticipate human errors in this system study. The case study will examine the operation of the ship's emergency shut down system during the LNG ship-to-ship bunkering procedure. The system is highly sophisticated, with a complicated operational system including a considerable number of workers and port authorities. Human error identification and risk models are developed independently, so HRA methodologies suited for quantifying human error should be identified. This section should also be addressed, as there is no referenced objective data on human error. Consideration should be given to ease of integration and compatibility with other methodologies.

5.4.4.2 Application of TOPSIS to determine optimal HRA for ESD system

Table 5. 18 summarises the evaluation results for the six HRA methods with scores ranging from 1 to 5 for this project's 14 evaluation criteria. According to the methods described in Section 5.3, a normalised evaluation matrix was derived in ***Table 5. 19***. ***Table 5. 20*** summarises the relative importance of each criterion for the ESD system during STS LNG bunkering procedures.

Table 5. 18 Evaluation matrix for LNG ESD system

Alternatives	Criteria													
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
BN-CREAM (A_1)	3	4	3	3	4	4	3	4	5	5	3	2	2	2
CREAM (A_2)	3	4	3	3	4	4	3	3	3	3	2	3	3	3
SPAR-H (A_3)	3	3	4	4	3	2	4	3	3	3	4	3	4	4
SLIM (A_4)	3	3	3	4	4	3	3	3	3	4	5	3	3	5
THERP (A_5)	4	2	4	2	2	3	3	3	1	1	1	4	2	2
HEART (A_6)	4	5	3	2	2	3	3	3	1	2	2	4	2	2

Table 5. 19 Normalised evaluation matrix for LNG ESD system

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.36380	0.45004	0.36380	0.39392	0.49614	0.50395	0.38411
CREAM	0.36380	0.45004	0.36380	0.39392	0.49614	0.50395	0.38411
SPAR-H	0.36380	0.33753	0.48507	0.52523	0.37210	0.25198	0.51215
SLIM	0.36380	0.33753	0.36380	0.52523	0.49614	0.37796	0.38411
THERP	0.48507	0.22502	0.48507	0.26261	0.24807	0.37796	0.38411
HEART	0.48507	0.56254	0.36380	0.26261	0.24807	0.37796	0.38411
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.51215	0.68041	0.62500	0.39057	0.25198	0.29488	0.25400
CREAM	0.38411	0.40825	0.37500	0.26038	0.37796	0.44233	0.38100
SPAR-H	0.38411	0.40825	0.37500	0.52076	0.37796	0.58977	0.50800
SLIM	0.38411	0.40825	0.50000	0.65094	0.37796	0.44233	0.63500
THERP	0.38411	0.13608	0.12500	0.13019	0.50395	0.29488	0.25400
HEART	0.38411	0.13608	0.25000	0.26038	0.50395	0.29488	0.25400

Table 5. 20 Weights and normalised weights for criteria for LNG ESD system

Criteria	C1	C2	C3	C4	C5	C6	C7
W_j	3	3	3	4	5	4	4
w_j	0.06	0.06	0.06	0.08	0.10	0.08	0.08
Criteria	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**
W_j	1	2	3	4	4	4	5
w_j	0.02	0.04	0.06	0.08	0.08	0.08	0.10

Table 5. 21 Weighted normalised evaluation matrix for LNG ESD system

Alternatives	Criteria						
	C1	C2	C3	C4	C5	C6	C7
BN-CREAM	0.02227	0.02755	0.02227	0.03216	0.05063	0.04114	0.03136
CREAM	0.02227	0.02755	0.02227	0.03216	0.05063	0.04114	0.03136
SPAR-H	0.02227	0.02066	0.02970	0.04288	0.03797	0.02057	0.04181
SLIM	0.02227	0.02066	0.02227	0.04288	0.05063	0.03085	0.03136
THERP	0.02970	0.01378	0.02970	0.02144	0.02531	0.03085	0.03136
HEART	0.02970	0.03444	0.02227	0.02144	0.02531	0.03085	0.03136
Alternatives	**Criteria**						
	C8	C9	C10	C11	C12	C13	C14
BN-CREAM	0.01045	0.02777	0.03827	0.03188	0.02057	0.02407	0.02592
CREAM	0.00784	0.01666	0.02296	0.02126	0.03085	0.03611	0.03888
SPAR-H	0.00784	0.01666	0.02296	0.04251	0.03085	0.04814	0.05184
SLIM	0.00784	0.01666	0.03061	0.05314	0.03085	0.03611	0.06480
THERP	0.00784	0.00555	0.00765	0.01063	0.04114	0.02407	0.02592
HEART	0.00784	0.00555	0.01531	0.02126	0.04114	0.02407	0.02592

Based on ***Table 5. 21***, the best and worst alternatives are obtained as follows.

The worst alternative A^w = (0.02227, 0.01378, 0.02227, 0.02144, 0.02531, 0.02057, 0.03136, 0.00784, 0.00555, 0.00765, 0.01063, 0.03085, 0.02407, 0.02592) and

The best alternative A^b = (0.02970, 0.03444, 0.02970, 0.04288, 0.05063, 0.04114, 0.04181, 0.01045, 0.02777,0.03827, 0.05314, 0.04114, 0.04814, 0.06480)

The distance (S_i^-) between the target alternative (A_i) and the worst condition (A^w) and the distance (S_i^+) between the alternative (A_i) and the best condition (A^b) were determined in

Table 5. **22**.

Table 5. 22 The distances of S_i^- and S_i^+ and the rank of alternatives for the LNG ESD system

Alternatives	S_i^-	S_i^+	Similarity	Rank
A_1	0.06	0.06	0.50	3
A_2	0.05	0.05	0.47	4
A_3	0.06	0.04	0.60	2
A_4	0.07	0.03	0.70	1
A_5	0.02	0.08	0.18	6
A_6	0.03	0.07	0.28	5

In conclusion, during the ship-to-ship LNG bunkering process research project, the SLIM approach was determined to be the most effective way for the emergency shutdown system (ESD).

5.5 Chapter summary

Individual four evaluations using 14 criteria were undertaken in this chapter to determine the optimal HRA approach for each case study, including emergency steering in Chapter 6, engine room fire drill in Chapter 7, man overboard drills in Chapter 8, and STS LNG bunkering process in Chapter 9. It was identified that there is no structured approach for selecting an appropriate HRA technique for a specific case study by considering multi criteria. This chapter demonstrated a process which will assist in systematically considering contextual factors and selecting the most feasible HRA technique for the case under investigation. The assessment result indicated that the most optimal HRA techniques are BN-CREAM for emergency steering, CREAM for engine room fire drills, SPAR-H for man overboard drills, and SLIM for the LNG ESD system, respectively. Each case study will be discussed in detail from Chapter 6 to Chapter 9. As previously stated, strategy in this chapter does not assess the relative superiority of the HRA method itself but rather determines the best appropriate HRA method for a particular study environment.

6 Determination of the human error mode

6.1 Chapter overview

This chapter introduces a method of determining the contextual mode for overall human error probability estimation using the BN-CREAM method. This BN-CREAM method aims to determine the need for more specific HRA research and to support urgent analysis when required by providing a simple and imminent calculation method. This chapter first explains the necessity and background of this methodology in Section 6.2. Then the specific details of the methodology are introduced in Section 6.3. Following this, a case study on the emergency steering operations using this methodology will be illustrated in Section 6.4. Finally, in section 6.5, findings and discussion will be presented.

6.2 Background of the research

Human operators play a significant role in the system and operational safety of the maritime industry. Numerous HRA studies have been done over the last few decades to improve our understanding of human roles in safety-critical sectors and to forecast the frequency of possible human errors and their effects. However, human reliability analysis is a time-consuming and resource-intensive exercise that involves a significant number of human resources, available data, and effort. As a result, adding the human aspect to risk assessments of emergency decision-making that may occur in dynamic circumstances is a real practical challenge. Additionally, there is no way to determine whether a detailed HRA study on the system or operation is necessary. As seen in ***Figure 6. 1***, a simple and reliable approach for rapidly providing information on human errors coincides with the requirement for further detailed HRA study on the system or operation. The basic CREAM method estimates total HEP by analysing context using PSFs without task analysis. The following studies offered straightforward approaches to addressing these research needs.

Fujita and Hollnagel (2004) introduced systematic procedures for calculating mean failure rates as a function of the CPC without making assumptions about individual human actions by establishing a simple mathematical manipulation. Konstandinidou et al. (2006) have developed a Fuzzy modelling system to estimate the probability of erroneous human action in specific industrial and working contexts based on CREAM methodology. The developed

Fuzzy logic consists of 9 input variables similar to CPCs and an If-Then knowledge-based Fuzzy inference system to predict a crisp value that is a failure probability of human operation. He et al. (2008) provided a simplified CREAM prospective quantification process to provide an easily practicable approach to getting the numeric results. The method can be applied to both the basic and extended methods. There have also been attempts to develop models that could directly estimate overall HEPs using BNs. Islam et al. (2018) introduced a BN model to estimate HEP using priority probability and CPT (conditional probability table) from expert groups. Their research determined the impact of internal and external factors on human performance with a case study for ship maintenance activities. The BN model provides flexible HEPs that could be assigned with new input variables. It made it possible to predict HEPs across various maritime scenarios dynamically. Despite its effectiveness on HEPs, the BN models are subject to the standardised contexts, thereby the same level of PSFs, against disparate activities. Furthermore, the direct inference logic model is weak to capture the significant difference among subtasks under a similar situation without considering different levels of task because contributing factors do not fully address the characteristic of the different levels of functions.

While the simplified CREAM method in the past provided an easily accessible process for obtaining human error probability mode, numerous assumptions were inevitably made due to the uncertainty raised from excessive simplification ideas. Specifically, in the Fuzzy logic model, the evaluation of individual PSFs is directly reflected in the output. As a result, it can perform fast human error calculations, but the limitation is that much information is lost by If-Then Rule and dependence between PSFs cannot be considered. On the other hand, the BN model considers interdependence among the PSFs. Still, to reduce computational load, the rating of individual PSF does not directly affect the model's output by grouping PSF. This process lowers the sensitivity of the model and further complicates the network. Therefore, this paper proposes a new BN-CREAM model to prevent the loss of information and strengthen the reflection of dependence and simplification.

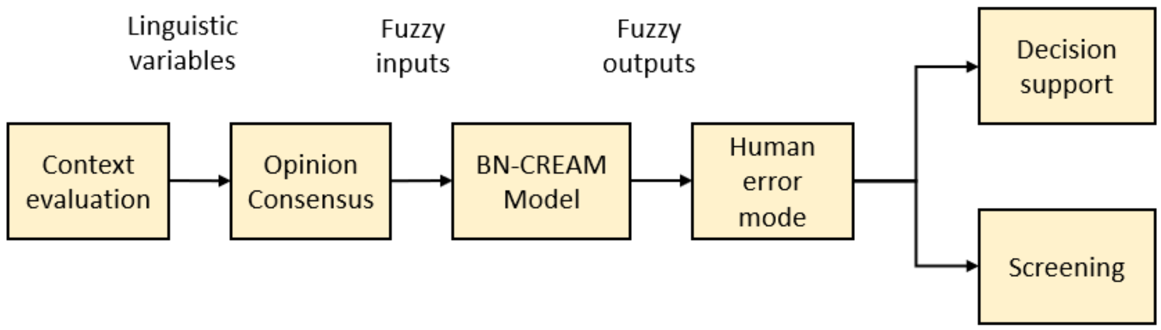

Figure 6. 1 Abstract for the simplified human error mode estimating model

6.3 Methodology

This section proposes a hybrid approach combining Fuzzy theory and Bayesian Network (BN) with CREAM to predict human error modes. The proposed BN-CREAM model provides maximum simplification and rapid human error modes. The proposed model is developed to predict the overall human error mode quickly and conveniently by analysing contextual factors without analysing individual tasks. The model is designed to directly reflect the ratings of individual PSFs in model output human error mode without grouping PSFs to avoid losing information and sensitivity. This model reflects PSF interdependence through BN and provides a more intuitive representation and easy modelling. Also, a Fuzzy multiple attributive group decision-making methodology by Ölçer and Odabaşi (2005) is employed and customised for the opinion aggregation to minimise the subjectivity of experts' judgment. The flow chart of the proposed approach is shown in ***Figure 6. 2***.

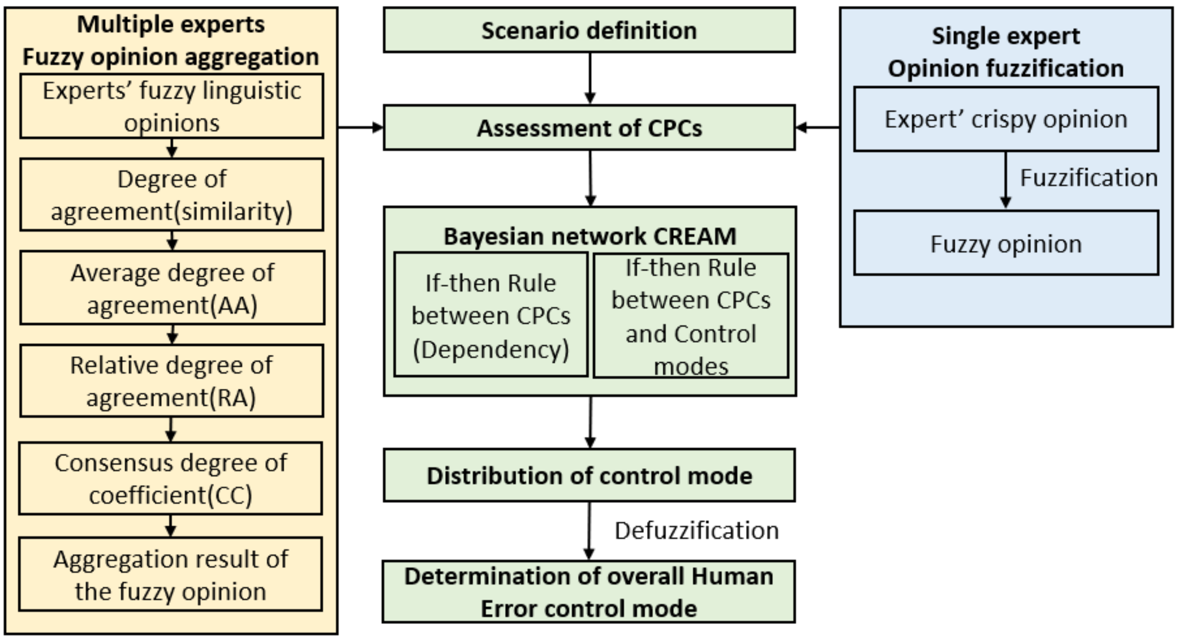

Figure 6. 2 Flowchart of the proposed approach

6.3.1 Common performance condition (CPC) Assessment

Once the definition of the scenario, which is the first step in analyses, is completed, an evaluation of individual CPCs is initiated.

6.3.1.1 Define CPCs and Fuzzy membership development

The nine CPCs characterising performance shaping factors that impact human performance are listed in ***Table 6. 1***. Individual CPCs contain linguistic variables that indicate the CPC level likely to have an adverse or positive influence on human performance. In the original CREAM, the only linguistic variable is determined with a 100% degree of conviction to assess the relevant CPC (Ahn and Kurt, 2020). However, a small set of linguistic variables is insufficient to capture the effect of CPC on human reliabilities in practice. To illustrate the influence of CPC more accurately, Fuzzy sets are used since they are the appropriate method for dealing with the ambiguity and vagueness inherent in human error detection problems (Akyuz, 2016). As seen in ***Figure 6. 3***, ***Figure 6. 4*** and ***Figure 6. 5***, each CPC is associated with three or more Fuzzy sets that explain the influence of the CPC. The Fuzzy trapezoidal number is used in this study, and the associated Fuzzy numbers for each CPC level are produced and listed in **Table 6. *2***.

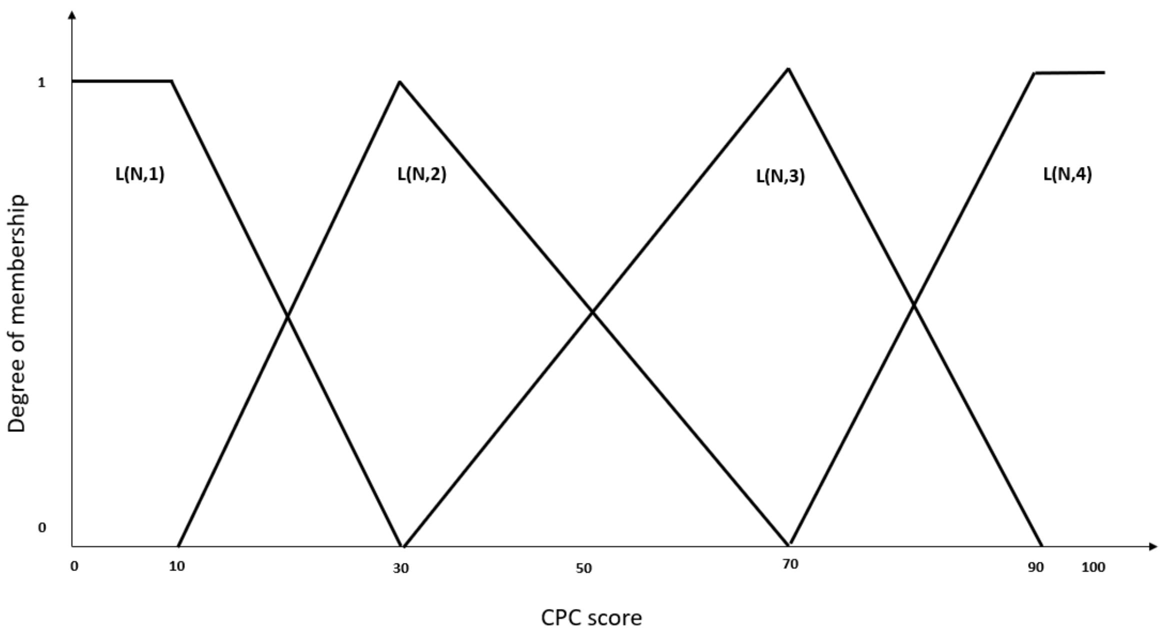

Figure 6. 3 Fuzzy membership for four linguistic scales

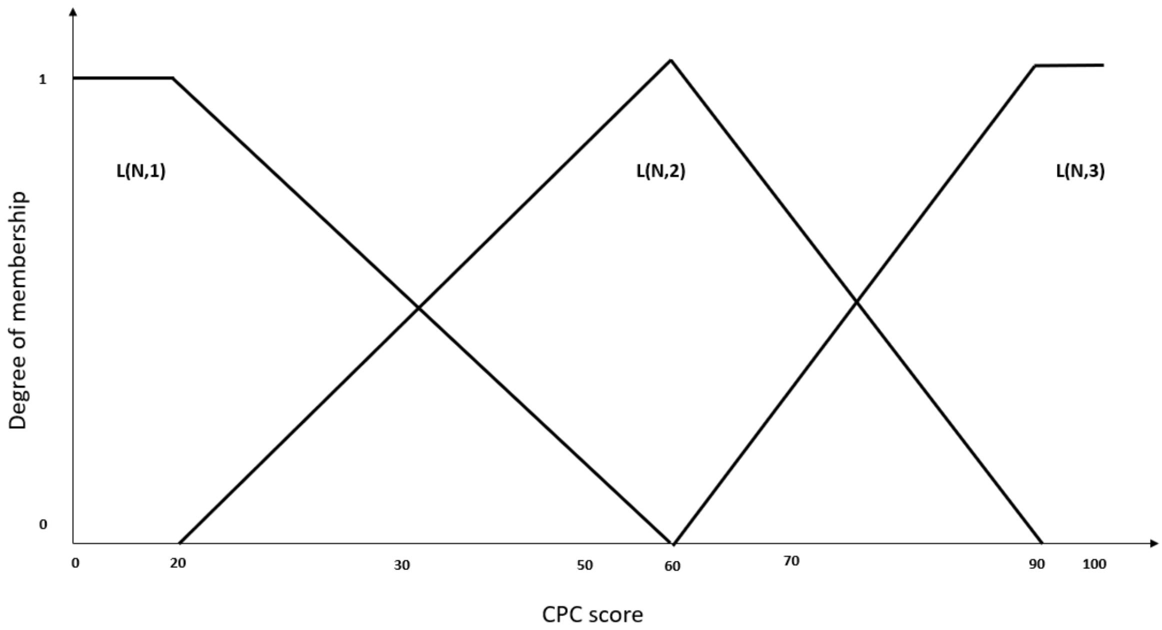

Figure 6. 4 Fuzzy membership for three linguistic scales

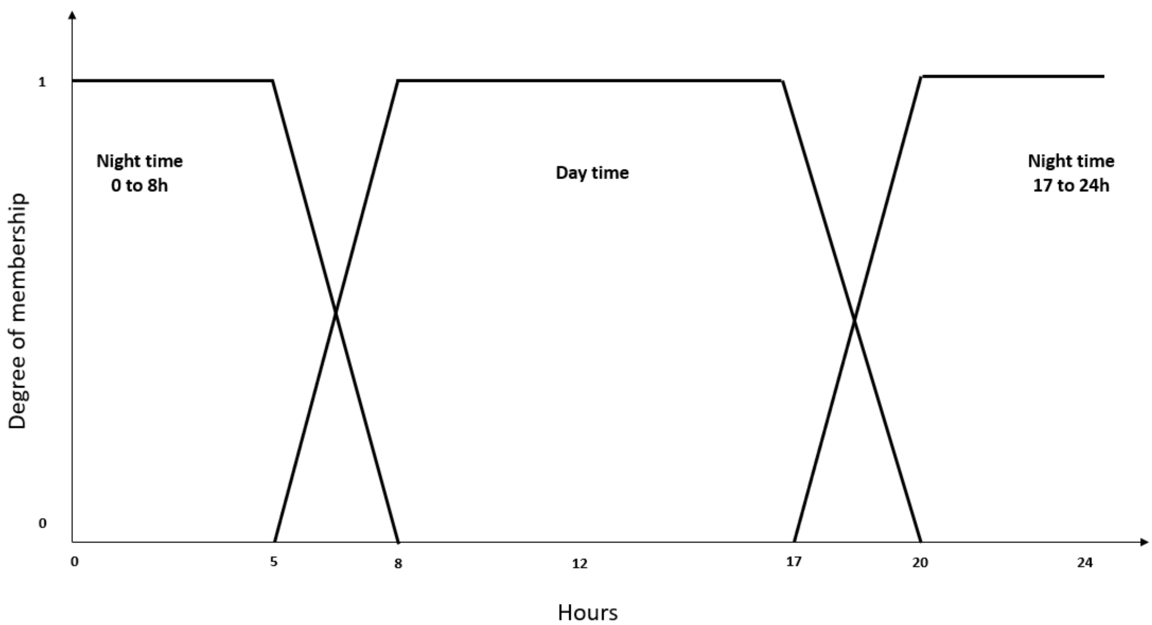

Figure 6. 5 Fuzzy membership for Time of day

Table 6. 1 Common Performance Condition(Hollnagel, 1998)

CPC name	Level description & Linguistic scales
Adequacy of organisation	The quality of the roles and responsibilities of team members, additional support, communication system, Safety Management System, instructions and guidelines for externally oriented activities, role of external agencies, etc. Very efficient / Efficient / Inefficient / Deficient
Working condition	The nature of the physical working conditions such as ambient lighting, glare on screens, noise from alarm, interruptions from the task, etc. Advantageous / Compatible / Incompatible
Adequacy of MMI and operational condition	The Man-Machine interface in general, including the information available on control panels, computerised workstations, and operational support provided by specially designed decision aids Supportive / Adequate / Tolerable / Inappropriate
Availability of procedures/plan	Procedures and plans include operating and emergency procedures, familiar patterns of response heuristics, routines, etc. Appropriate / Acceptable / Inappropriate
Number of simultaneous goals	The number of tasks a person is required to pursue or attend to simultaneously (i.e., evaluating the effects of actions, sampling new information, assessing multiple goals etc.) Fewer than capacity / Matching current capacity / More than capacity
Available time	The time available to carry out a task corresponds to how well the task execution is synchronised to the process dynamics Appropriate / Temporarily inadequate / Continuously inadequate
Time of day (Circadian rhythm)	The time of day (or night) describes the time at which the task is carried out, particularly whether or not the person is adjusted to the current time (circadian rhythm). Typical examples are the effects of shift work. It is a well-established fact that the time of day influences the quality of work and that performance is less efficient if the normal circadian rhythm is disrupted Daytime (adjusted) / Night-time (unadjusted)
Adequacy of training and experience	The level and quality of training provided to operators as feminisation to new technology, refreshing old skills, etc. It also refers to the level of operational experience. Adequate, high experience / Adequate, limited experience / Inadequate
Crew collaboration quality	The quality of the collaboration between crew members, including the overlap between the official and unofficial structure, the level of trust, and the general social climate among crew members Very efficient / Efficient / Inefficient / Deficient

Table 6. 2 CPCs and Performance reliability with Fuzzy sets(Ahn and Kurt, 2020)

CPC name	CPC level (L_{ij})	The expected effect on performance	Fuzzy sets
CPC_1 Adequacy of organisation	Very efficient ($L_{1,4}$)	Improved	(70, 90, 100, 100)
	Efficient($L_{1,3}$)	Not significant	(30, 70, 70, 90)
	Inefficient($L_{1,2}$)	Reduced	(10, 30, 30, 70)
	Deficient($L_{1,1}$)	Reduced	(0, 0, 10, 30)
CPC_2 Working condition	Advantageous ($L_{2,3}$)	Improved	(60, 90, 100, 100)
	Compatible ($L_{2,2}$)	Not significant	(20, 60, 60, 90)
	Incompatible($L_{2,1}$)	Reduced	(0, 0, 20, 60)
CPC_3 Adequacy of MMI and operational condition	Supportive ($L_{3,4}$)	Improved	(70, 90, 100, 100)
	Adequate ($L_{3,3}$)	Not significant	(30, 70, 70, 90)
	Tolerable ($L_{3,2}$)	Not significant	(10, 30, 30, 70)
	Inappropriate ($L_{3,1}$)	Reduced	(0, 0, 10, 30)
CPC_4 Availability of procedures/plan	Appropriate($L_{4,3}$)	Improved	(60, 90, 100, 100)
	Acceptable ($L_{4,2}$)	Not significant	(20, 60, 60, 90)
	Inappropriate ($L_{4,1}$)	Reduced	(0, 0, 20, 60)
CPC_5 Number of simultaneous goals	Fewer than capacity ($L_{5,3}$)	Not significant	(60, 90, 100, 100)
	Matching current capacity ($L_{5,2}$)	Not significant	(20, 60, 60, 90)
	More than capacity ($L_{5,1}$)	Reduced	(0, 0, 20, 60)
CPC_6 Available time	Appropriate ($L_{6,3}$)	Improved	(60, 90, 100, 100)
	Temporarily inadequate	Not significant	(20, 60, 60, 90)
	Continuously inadequate	Reduced	(0, 0, 20, 60)
CPC_7 Time of day	Day-time 8h to 17h ($L_{7,3}$)	Not significant	(5, 8, 17, 20)
	Night-time 0h to 8h ($L_{7,2}$)	Reduced	(0, 0, 5, 8)
	Night-time 17h to 24h ($L_{7,1}$)	Reduced	(17, 20, 24, 24)
CPC_8 Adequacy of training and experience	Adequate, high experience ($L_{8,3}$)	Improved	(60, 90, 100, 100)
	Adequate, limited experience ($L_{8,2}$)	Not significant	(20, 60, 60, 90)
	Inadequate ($L_{8,1}$)	Reduced	(0, 0, 20, 60)
CPC_9 Crew collaboration quality	Very efficient ($L_{9,4}$)	Improved	(70, 90, 100, 100)
	Efficient ($L_{9,3}$)	Not significant	(30, 70, 70, 90)
	Inefficient ($L_{9,2}$)	Not significant	(10, 30, 30, 70)
	Deficient ($L_{9,1}$)	Reduced	(0, 0, 10, 30)

6.3.1.2 Selection of experts

The selection of experts for CPC evaluation is a critical process. In this study, a number of experts can be flexibly selected from a single expert to multiple experts group according to research resources such as human resources, time, and needs. Evaluation by a single expert is not recommended if resources are allowed. Still, to cope with dynamic situations, the method by a single expert was developed to provide a simple process and practical way for the industry. The evaluation methods by a single expert and a group of experts have different approaches.

6.3.1.3 CPC evaluation by a single expert

To evaluate individual CPCs, the expert is given scenarios, details, and questionnaires. The expert is asked to rate specific CPCs on a scale of 0 to 100. For example, the linguistic variables 'Inappropriate', 'Tolerable', 'Adequate', and 'Supportive' are used to assess the 'Adequacy of MMI and operational condition'. The horizontal axis indicates the numerical score of this CPC, which ranges from 0 to 100, with 0 being the most negative value and 100 being the most positive. In contrast, the vertical axis reflects the degree of membership, which ranges from 0 to 1. If the CPC3 score is assigned as 80 points, the single numerical score can be fuzzified to represent a distribution of beliefs with 75% trust in Supportive(L3,4) and 25% faith in Adequate(L3,3) as shown in ***Figure 6. 6***.

Trapezoidal Fuzzy set expressed as (a,b,c,d) and membership function μ_{ij} for random score x is obtained through fuzzification.

$$\mu_{ij} = \begin{cases} \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b \le x \le c \qquad \text{where } a \le b \le c \le d \\ \frac{d-x}{d-c}, & c \le x \le d \\ 0, & \text{Otherwise} \end{cases} \qquad (6\text{-}1)$$

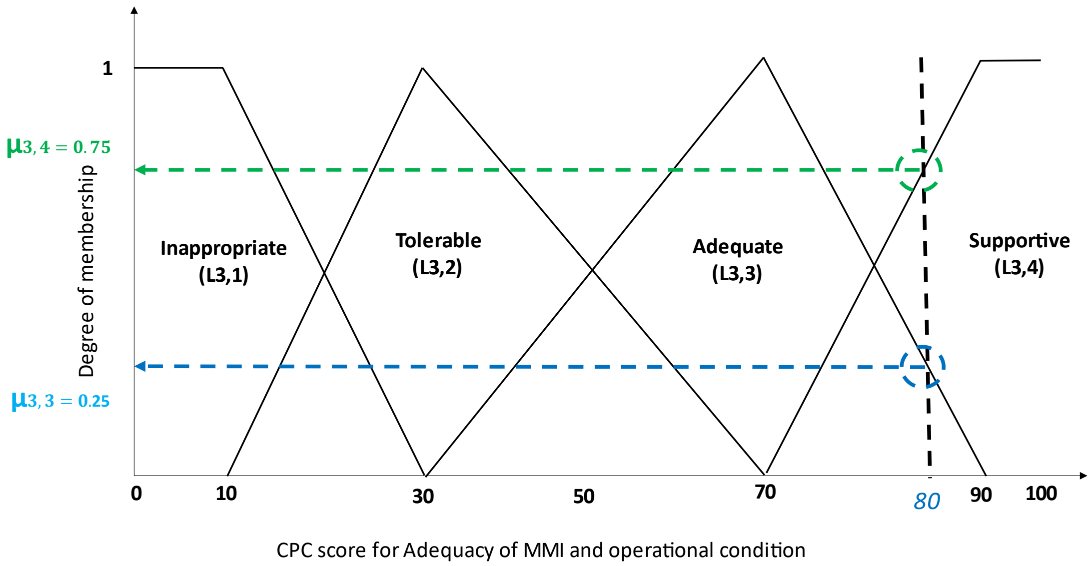

Figure 6. 6 Membership functions for Adequacy of MMI and operational condition

6.3.1.4 CPC evaluation by multiple experts

Experts are required to evaluate each CPC using the appropriate linguistic variables. Linguistic scale for CPC level and their corresponding Fuzzy set developed and provided in **Table 6. *2***. The purpose of applying the Fuzzy opinion aggregation is to translate the experts' multiple qualitative assessments of CPC score into a single aggregated opinion with Fuzzy opinion, then convert it into a crisp value through defuzzification. Once experts are selected, their relative importance needs to be considered as a heterogeneous group depending on their background. An example of experts' evaluation criteria is given in **Table 6. 3**. The normalised weighted importance of experts is utilised to make an opinion consensus. This chapter does not discuss the technique for aggregating opinions because it is detailed in Section 7.2.1.1. Once aggregated scores are derived, fuzzification for each CPCs is required to convert the score into a distribution of beliefs for linguistic scales.

Table 6. 3 Example of the experts' evaluation criteria

Constitution	Classification	Ex1	Ex2	Ex3	Ex4	Ex5	Score
Professional position	Senior Academic			5			5
	Junior Academic						4
	Class. Surveyor or Government inspector or ISM auditor (senior level)	5			5	5	5
	Class. Surveyor or Government inspector or ISM auditor (Junior level)		4				4
	Senior Mariner						5
	Junior Mariner						3
	Ship Superintendent (Senior)						5
	Ship superintendent (Junior)						4
	Others						between 1 and 5
Total Service time	30 years and above						5
	20 - 29					4	4
	between 10 and 19	3	3	3	3		3
	between 6 and 9						2
	Below 5						1
Education level	PhD			5			5
	Master	4	4		4	4	4
	Bachelor						3
	Diploma						2
	High school						1
Fire drill experience as a crew or ISM auditor	Yes	3	3	3	3	3	3
	No						0
Total		15	14	16	15	16	76
Normalised weighted importance		**0.2**	**0.18**	**0.21**	**0.2**	**0.21**	**1**

6.3.2 Bayesian network modelling

Once an expert assigns a rating of CPCs, the evaluation value of the nine individual CPC becomes the input value of the BN-CREAM model. The main aims of the BN-CREAM model are to adjust the dependency of CPCs and present the distribution of human error control mode to determine human error control mode as shown in ***Figure 6. 7***. Then, the belief distribution for the control mode is converted to a form of human error probability by defuzzification.

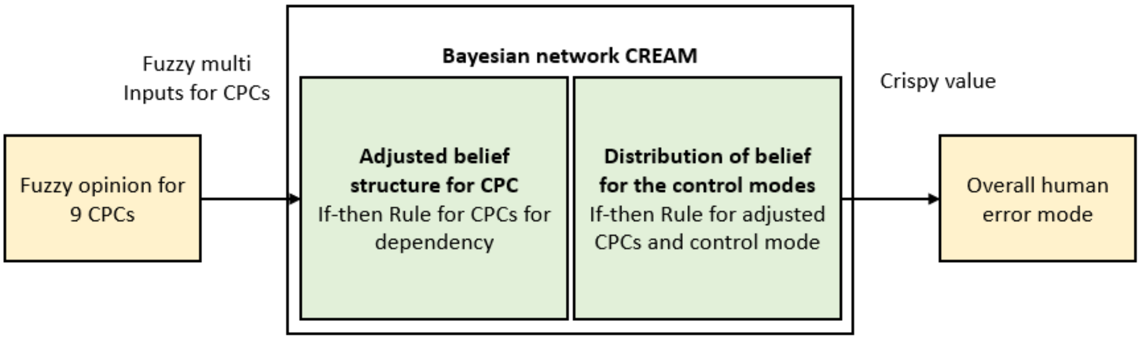

Figure 6. 7 BN-CREAM model

6.3.2.1 Adjusted belief structure for CPC

In the preceding stage, each CPC is expressed through fuzzification as a belief structure in single and multiple expert judgement cases. However, the dependency between CPCs should be recognised, and CPCs should be changed accordingly, as CPCs are not entirely independent of the influence of other CPCs. Thus, it needs to define the principles governing the reciprocal impacts of CPCs. For example, the rule of the 4th row indicates that 'Crew collaboration quality' depends on both 'adequacy of organisation' and 'adequacy of training and experience'. If 'crew collaboration of quality' is inefficient (Neutral) AND 'Adequacy of organisation' is very efficient (Positive) AND 'Adequacy of training and experience' is Adequate, high experience (Positive), then "Crew collaboration quality is adjusted to positive from neutral. Interactive relations can be modelled by a Bayesian network technique (Yang et al., 2013) and enable presenting rather complex systems (Hänninen, 2014). Bayesian network model based on rules acquires four new adjusted CPCs from the nine original CPCs. Adjusted CPCs are also represented by a new belief structure as follows.

$CPC_i' = ((\mu_{i1}', L_{i1}), (\mu_{i2}', L_{i2}), (\mu_{i3}', L_{i3}), (\mu_{ij}', L_{ij}))$, where i= [1, 9] and j=[1, 4] (6-2)

Fuzzy sets for nine CPCs are entered into a model as an input variable, and 4 CPCs are adjusted based on rules of dependency, as shown in ***Figure 6. 8***.

Table 6. 4 Rules for adjusting CPCs (Hollnagel,1998)

Adjusted CPC	Affecting CPCs				
Working Conditions (4/5)	Adequacy of organisation	Adequacy of MMI and operational support	Available time	Time of day	Adequacy of training and experience
Number of simultaneous goals (2/3)	Working Conditions	Adequacy of MMI and operational support	Availability of procedures and plans	-	-
Available time (4/5)	Working Conditions	Adequacy of MMI and operational support	Availability of procedures and plans	Number of simultaneous goals	Time of day
Crew collaboration quality (2/2)	Adequacy of organisation	Adequacy of training and experience	-	-	-

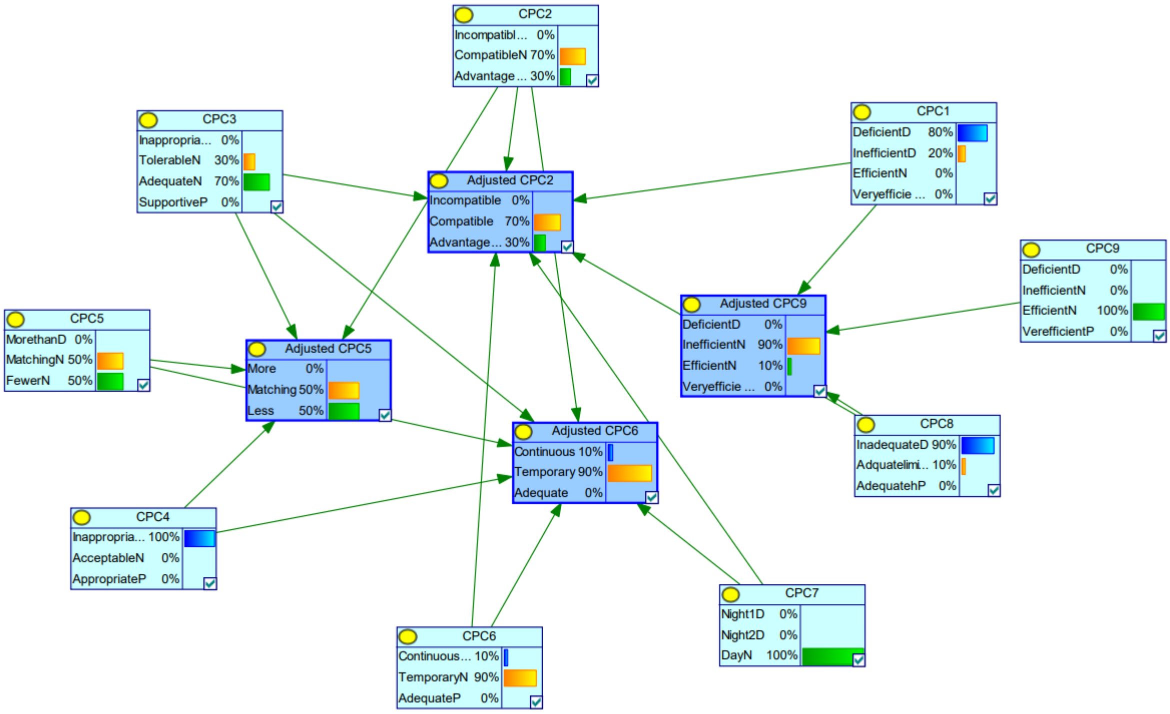

Figure 6. 8 Bayesian networks for dependency adjustment

6.3.2.2 Distribution of belief for control modes

Once CPCs are adjusted, rules between CPCs and Control mode need to be established to make a model. In the total 9 CPCs, 6 CPCs have three levels, and the remaining 3 CPCs have four. Therefore, the number of cases can produce 46,656 conditional probabilities to build a network, as shown in ***Table 6. 5***. This looks like a challenging task but is feasible in practice. Therefore, reducing conditional probabilities is not required by grouping CPCs that cause loss of information. The team has developed 46,656 rules based on conditional probabilities in ***Table 6. 6*** and **Table 6. *7***, considering the multiple Fuzzy sets of each input parameter and using the logical IF-AND operation. Each row in ***Table 6. 6*** shows each CPC's effect and their negative and positive effect sum. Once the negative and positive impact sum is calculated, the control mode is assigned based on ***Figure 6. 9***. For example, the If-Then Rule can be expressed below.

"IF L1,3 AND L2,1 AND L3,1 AND L4,1 AND L5,1 AND L6,1 AND L7,1 AND L8,1 AND L9,1, THEN the probability of $COCOM_j$ (j = 1, 2, 3, 4) is (0,0,0,1)."

Thus, the BN-CREAM model for COCOM and CPCs is built as shown in ***Figure 6. 10***.

Table 6. 5 List of the *conditional probabilities of* CPC

Number of cases	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
1	L1,4	L2,3	L3,4	L4,3	L5,3	L6,3	L7,3	L8,3	L9,4
2	L1,4	L2,3	L3,4	L4,3	L5,3	L6,3	L7,3	L8,3	L9,3
3	L1,4	L2,3	L3,4	L4,3	L5,3	L6,3	L7,3	L8,3	L9,2
4	L1,4	L2,3	L3,4	L4,3	L5,3	L6,3	L7,3	L8,3	L9,1
…..	….	….	….	….	….	….	….	….	….
46,653	L1,1	L2,1	L3,1	L4,1	L5,1	L6,1	L7,1	L8,1	L9,4
46,654	L1,1	L2,1	L3,1	L4,1	L5,1	L6,1	L7,1	L8,1	L9,3
46,655	L1,1	L2,1	L3,1	L4,1	L5,1	L6,1	L7,1	L8,1	L9,2
46,656	L1,1	L2,1	L3,1	L4,1	L5,1	L6,1	L7,1	L8,1	L9,1

Table 6. 6 Sum of CPC effect per conditional probability

Number of cases	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9	Sum of negative effect	Sum of positive effect
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	9	0
2	-1	-1	-1	-1	-1	-1	-1	-1	0	8	0
3	-1	-1	-1	-1	-1	-1	-1	-1	0	8	0
4	-1	-1	-1	-1	-1	-1	-1	-1	1	8	1
.....	....	....	....	....	....	....	....	....	....	....	....
46,653	1	1	1	1	0	1	0	1	-1	1	6
46,654	1	1	1	1	0	1	0	1	0	0	6
46,655	1	1	1	1	0	1	0	1	0	0	6
46,656	1	1	1	1	0	1	0	1	1	0	7

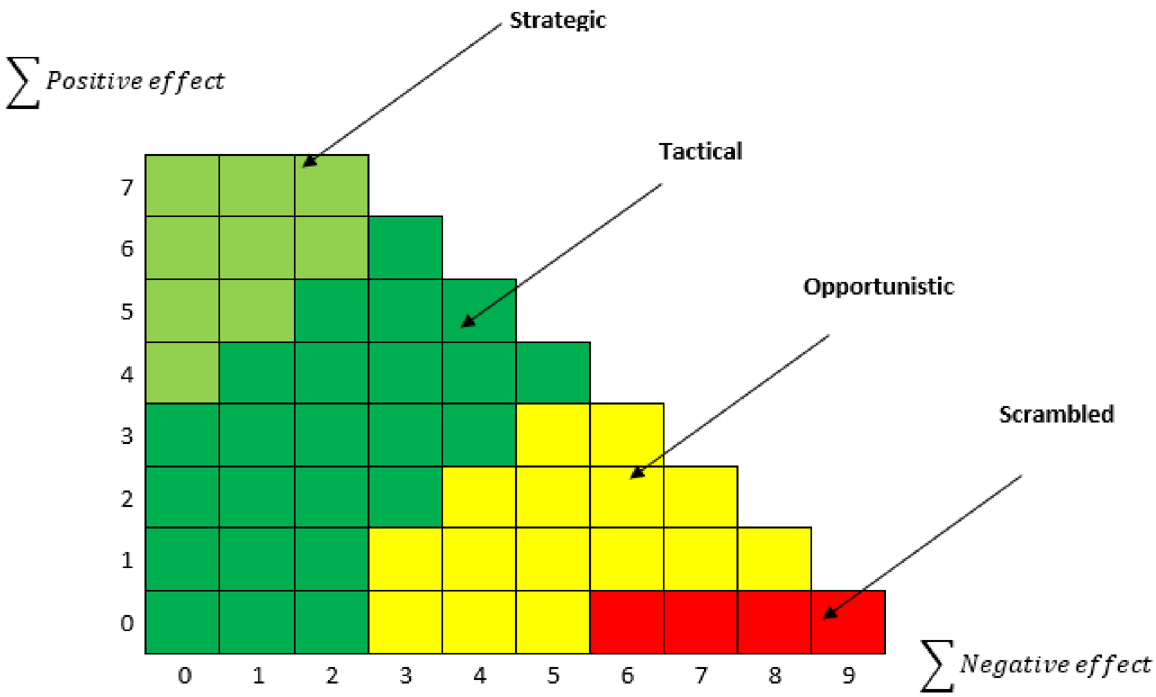

Figure 6. 9 Diagram of the Contextual Control Mode (Adopted from Hollnagel (1998))

Table 6. 7 The conditional probability table corresponding to the COCOM mode

Number of cases	Sum of the negative effect	Sum of the positive effect	COCOMs			
			Strategic	Tactical	Opportunistic	Scrambled
1	0	7	1	0	0	0
2	0	6	1	0	0	0
3	0	5	1	0	0	0
…	…	…	…	…	…	…
50	7	0	0	0	0	1
51	8	0	0	0	0	1
52	9	0	0	0	0	1

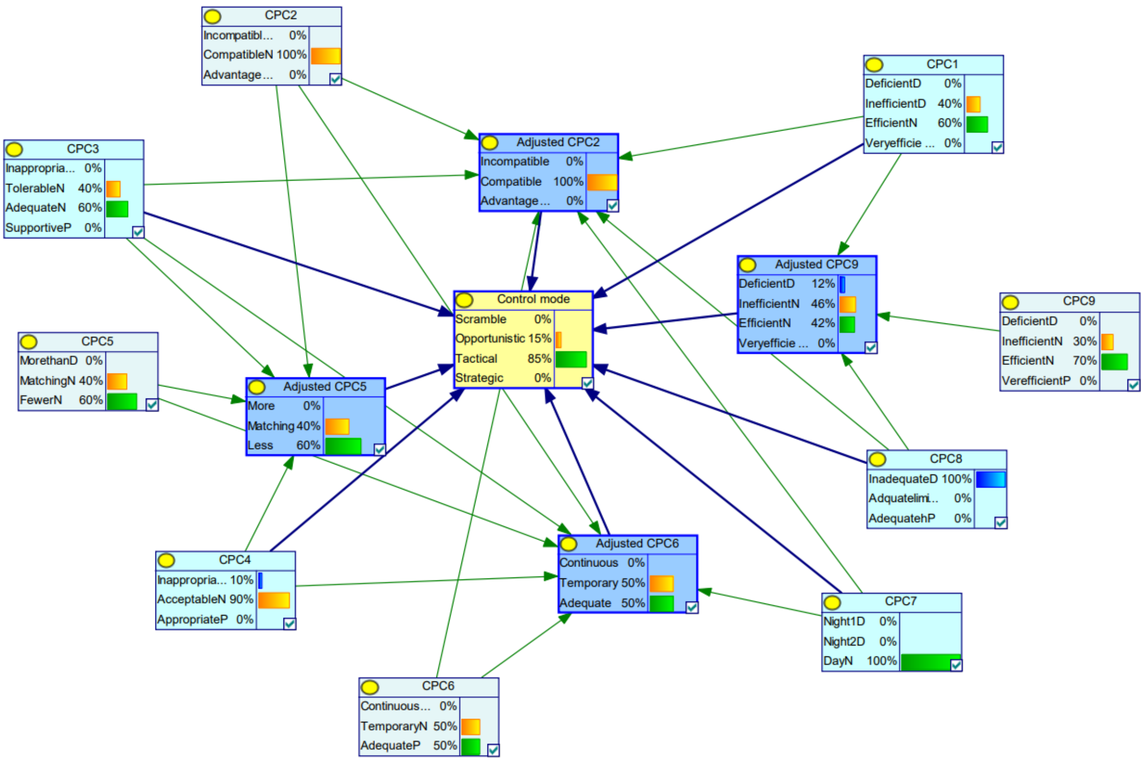

Figure 6. 10 Bayesian network for COCOM and CPCs

6.3.3 Defuzzification and human error probability

Defuzzification is a process of converting a Fuzzy conclusion to a crisp value. The Weighted Mean of Maxima (WMoM) is selected for this defuzzification. A set of belief degrees to the four control modes from the BN CREAM model is defuzzified into a crisp value as follows.

Crisp value (CV) = $\sum_{i=1}^{4} A_i^k * w_i$ (6-3)

Where w_i is the significant value of the i-th Fuzzy membership function.

The weighted value of a Fuzzy membership function is abscissa when the Fuzzy membership function is a maximum value. Membership functions have been developed based on the human failure probability interval in

Table 6. 8, as shown in ***Figure 6. 11***. The value of significance w_i can be calculated as -3.651, -2, -1.151 and -0.5. The final step is to convert a crisp value to human error probability since the CV is a logarithm value of human failure probability as below.

HEP (human error probability) = 10^{CV} (6-4)

Table 6. 8 Control mode and action failure probability (Hollnagel, 1998)

Control mode	Action failure probability
Strategic	0.5E-5< p < 1.0E-2
Tactical	1.0E-3< p < 1.0E-1
Opportunistic	1.0E-2< p < 0.5E-0
Scrambled	1.0E-1< p < 1.0E-0

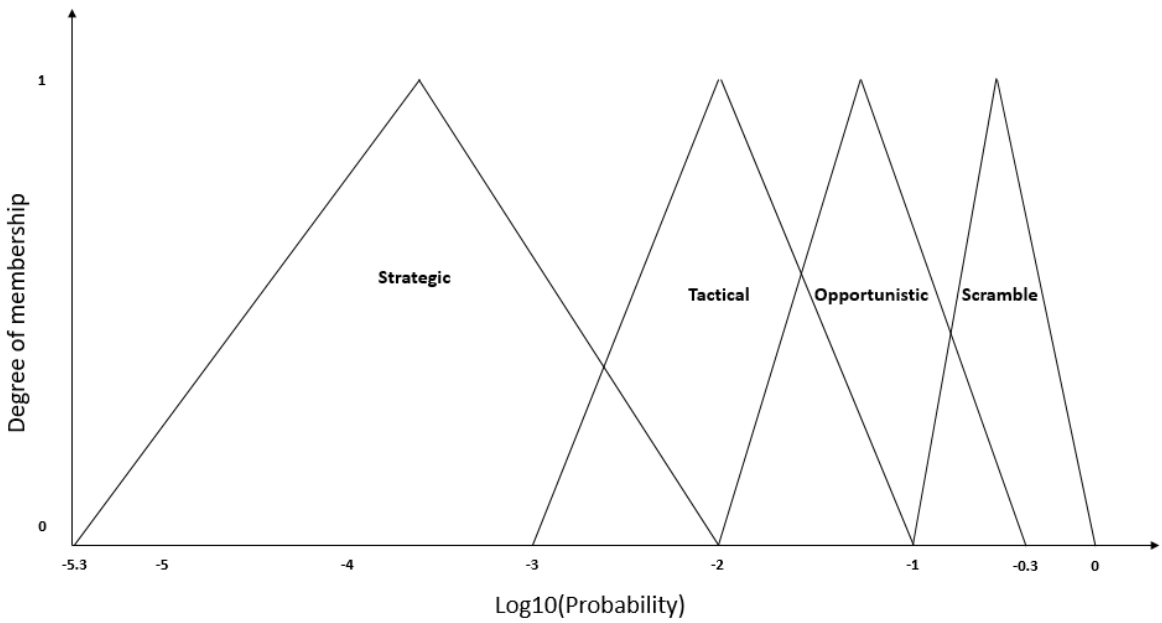

Figure 6. 11 Membership functions for control modes

6.4 Application of the Bayesian network CREAM to emergency steering operation (Case study 1)

The emergency steering had been chosen to demonstrate the proposed approach. An emergency steering system is deployed when the ship's primary steering system fails. Thus, emergency steering at sea is a life-threatening circumstance in which crews must effectively accomplish emergency steering duties with limited resources such as human resources, equipment, and time. The emergency steering system varies according to ship type, size, and year of construction. As a result, emergency steering methods depend on the ship's steering system. For this case study, the specific procedure for emergency steering gear operation was chosen and specified in Section 6.4.1. Section 6.4.2 describes an emergency steering operation to analyse contextual factors and anticipate total HEP without regard for specific human activity.

6.4.1 Emergency steering procedure and task analysis

The conceptual diagram for emergency steering for this case study is defined as shown in ***Figure 6. 12***, and procedures are described below.

- Both crews in the steering gear room and on the bridge should familiarise themselves with the process and diagram for emergency steering, which should be displayed in the steering gear room and on the bridge.
- Both crews maintain good communication for emergency operations via ship's telephone systems or two-way VHF.
- The crew in the steering gear room switch off the power supply from the panel of the steering gear system.
- The crew in the steering gear room change the mode of operation by selecting the switch for the motor which is supplied with emergency power.
- On the bridge, crew members remove a safety pin from the manual helms wheel. (Note that regular manual operation is always in the cut-off mode during normal operation.)

- The crew on the bridge order for the rudder angle.
- Crew members in the steering gear room should immediately turn the wheel and check the rudder angle indicator upon receiving commands.

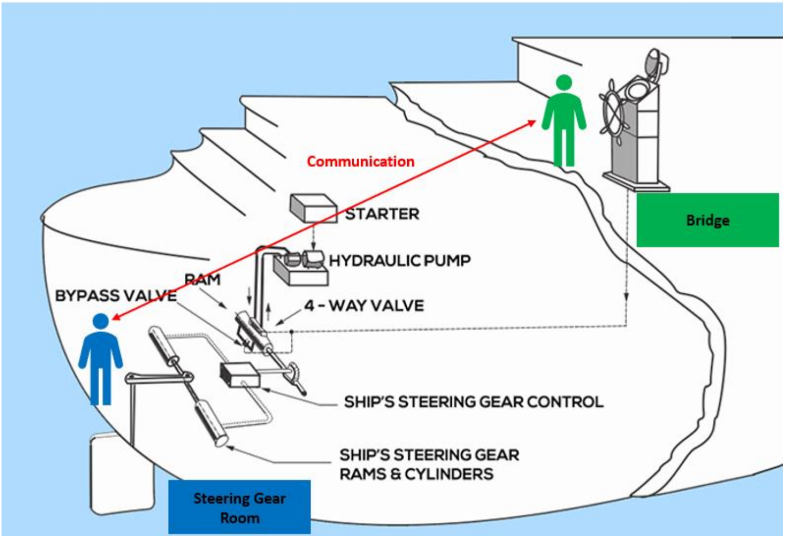

Figure 6. 12 Conceptual diagrams for an emergency steering

(Adopted from Marine-insight (2021))

6.4.2 Scenario description

The emergency steering scenario is detailed to illustrate the proposed method, with a particular emphasis on CPCs for evaluation.

The ship's main steering gear system was broken down at 2 pm while sailing through the ocean. Accordingly, the captain ordered an emergency steering operation, two crew members of emergency personnel moved to the steering room, and the captain and the wheeler were in charge of manoeuvring the ship on the bridge. For communication between the steering gear room and the bridge, onboard phones were used, but two-way VHFs were also placed in case of an emergency. The steering gear room was very noisy, making it challenging to communicate. The temperature was 38°C, and the humidity was 70% in the steering gear room. The sea conditions and wind were generally good. The ship is a 20-year-old general cargo ship G/T 10,000, vessel's overall condition is deteriorating. The ship management company manages 50 ships, including both DOC and SMC certificates for individual ships in effect under the International Safety Management Code (ISM). Still, it has not obtained ISO certification for the quality management system. As a result of conducting an internal audit of the ship last month, five out of 10 confirmed nonconformities have yet to be rectified. A total of 17 crew members were on board, consisting of three nationalities. Three crew members were replaced when they came to the main voyage. The emergency steering system of this ship shall be adjusted in the steering room according to the instructions of the bridge after changing the mode in the steering room. In preparation for such an accident, the shipping company provided a list of duties and responsibilities to the ship, while emergency control procedures and diagrams were provided in the steering gear room and the bridge. Recently an emergency steering training was conducted two months ago.

6.3.3 Common performance condition assessment

To demonstrate the proposed method, evaluations were undertaken separately by a single expert in section 6.3.3.1 and by a group of experts in section 6.3.3.2.

6.3.3.1 BN-CREAM modelling by a single expert

For assessment, a single expert was asked to assign CPC scores, and judgement is as shown in ***Table 6. 9***. Rated scores for each CPC were fuzzified into the distribution of belief for linguistic level as shown in ***Table 6. 10***. Fuzzified scores were entered into the BN-CREAM model as inputs, as shown in ***Figure 6. 13***, which was illustrated by Genie software. The output offers 58% for Opportunistic and 42% for Tactical modes. The overall human error probability for the given study was 3.11E-02 by defuzzifying the belief distribution for COCOM.

Table 6. 9 Single expert's evaluations of CPCs

CPC	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
SCORE	50	30	70	70	60	40	14h	70	65

Table 6. 10 Fuzzification and their distribution of belief for linguistic scales (Single expert)

Linguistic level	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
L(i,1)	0	0.75	0	0	0	0.5	0	0	0
L(i,2)	0.5	0.25	0	0.67	1	0.5	1	0.67	0.125
L(i,3)	0.5	0	1	0.33	0	0	0	0.33	0.875
L(i,4)	0	N/A	0	N/A	N/A	N/A	N/A	N/A	0

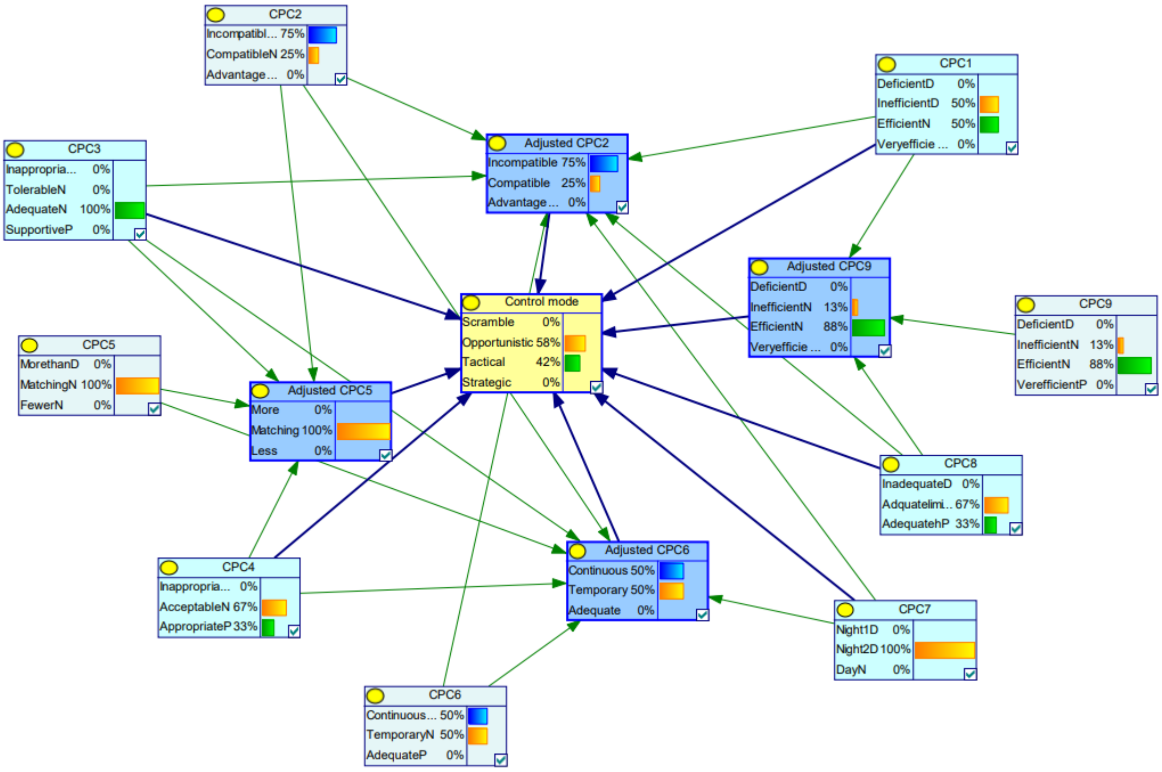

Figure 6. 13 BN-CREAM by single expert judgement

6.3.3.2 BN-CREAM modelling by multiple experts

For evaluation of the contextual factors, three experts are asked to assign CPC scores with linguistic scales, as shown in ***Table 6. 11***. In a group of experts' judgement, the relative importance among experts is considered as a heterogeneous group depending on their background by expert weighting criteria in ***Table 6. 3***. Experts' relative importance was assigned as 0.33,0.31 and 0.36 for this analysis. For group experts' judgement, opinion aggregation from CPC_1 to CPC_9 were conducted. A relaxation factor β is assumed to be 0.5. For example, a specific aggregation for CPC1 is illustrated in ***Table 6. 12***. The aggregated Fuzzy opinions are defuzzified and listed in ***Table 6. 13***. Once experts' judgement and Fuzzy opinion aggregation are completed, the next step is to convert the defuzzified CPC scores to Fuzzy membership again for BN-CREAM modelling input. The aggregated scores for each CPC were fuzzified into the distribution of belief for linguistic level, as shown in ***Table 6. 14***. Fuzzified scores were utilised to build the BN-CREAM model as inputs, as shown in ***Figure 6. 14***. The output shows 69% for Opportunistic and 31% for Tactical modes. The overall human error probability for the given study was 3.85E-02 by defuzzifying the distribution of belief for COCOM.

Table 6. 11 Group of experts' evaluations of CPCs

CPC	E1	E2	E3
CPC1	Efficient($L_{1,3}$)	Inefficient($L_{1,2}$)	Efficient($L_{1,3}$)
CPC2	Incompatible($L_{2,1}$)	Incompatible($L_{2,1}$)	Compatible ($L_{2,2}$)
CPC3	Supportive ($L_{3,4}$)	Adequate ($L_{3,3}$)	Adequate ($L_{3,3}$)
CPC4	Appropriate($L_{4,3}$)	Appropriate($L_{4,3}$)	Acceptable ($L_{4,2}$)
CPC5	More than capacity ($L_{5,1}$)	Matching current capacity ($L_{5,2}$)	Matching current capacity ($L_{5,2}$)
CPC6	Continuously inadequate (L6,1)	Temporarily inadequate (L6,2)	Continuously inadequate (L6,1)
CPC7	14h	14h	14h
CPC8	Adequate, limited experience (L8,2)	Adequate, high experience ($L_{8,3}$)	Adequate, limited experience (L8,2)
CPC9	Inefficient ($L_{9,2}$)	Inefficient ($L_{9,2}$)	Efficient ($L_{9,3}$)

Table 6. 12 Opinion aggregation under the CPC1

Standardised Fuzzy set (a1, a2, a3,a4)					
Expert	Judgement	a1	a2	a3	a4
Ex1	$L_{(1,3)}$	0.3	0.7	0.7	0.9
Ex2	$L_{(1,2)}$	0.1	0.3	0.3	0.7
Ex3	$L_{(1,3)}$	0.3	0.7	0.7	0.9
Degree of agreement(S)		Average degree of agreement (AA)			
S12	0.7	AA(Ex1)			0.85
S23	0.7	AA(Ex2)			0.7
S31	1	AA(Ex3)			0.85
Relative degree of agreement (RA)		Consensus degree coefficient (CC)			
RA(Ex1)	0.354167	CC(Ex1)			0.34208
RA(Ex2)	0.291667	CC(Ex2)			0.30083
RA(Ex3)	0.354167	CC(Ex3)			0.35708
Aggregated result		Aggregated Fuzzy set(S1,S2,S3,S4)			
		S1	S2	S3	S4
Rag(HT)		0.239833	0.579667	0.579667	0.83983
Defuzzification					0.55311
Normalised score					**55.3111**

Table 6. 13 Aggregated CPC scores

CPC scores	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
Aggregated	0.55	0.33	0.71	0.76	0.46	0.32	14h	0.65	0.45
Normalised	55.31	33.25	71.40	76.23	46.46	32.36	14h	65.32	45.36

Table 6. 14 Fuzzification and their distribution of belief for linguistic scales (Multiple experts)

Linguistic level	CPC1	CPC2	CPC3	CPC4	CPC5	CPC6	CPC7	CPC8	CPC9
L(i,1)	0	0.675	0	0	0.35	0.7	0	0	0
L(i,2)	0.375	0.325	0	0.467	0.65	0.3	1	0.833	0.625
L(i,3)	0.625	0	0.95	0.533	0	0	0	0.167	0.375
L(i,4)	0	N/A	0.05	N/A	N/A	N/A	N/A	N/A	0

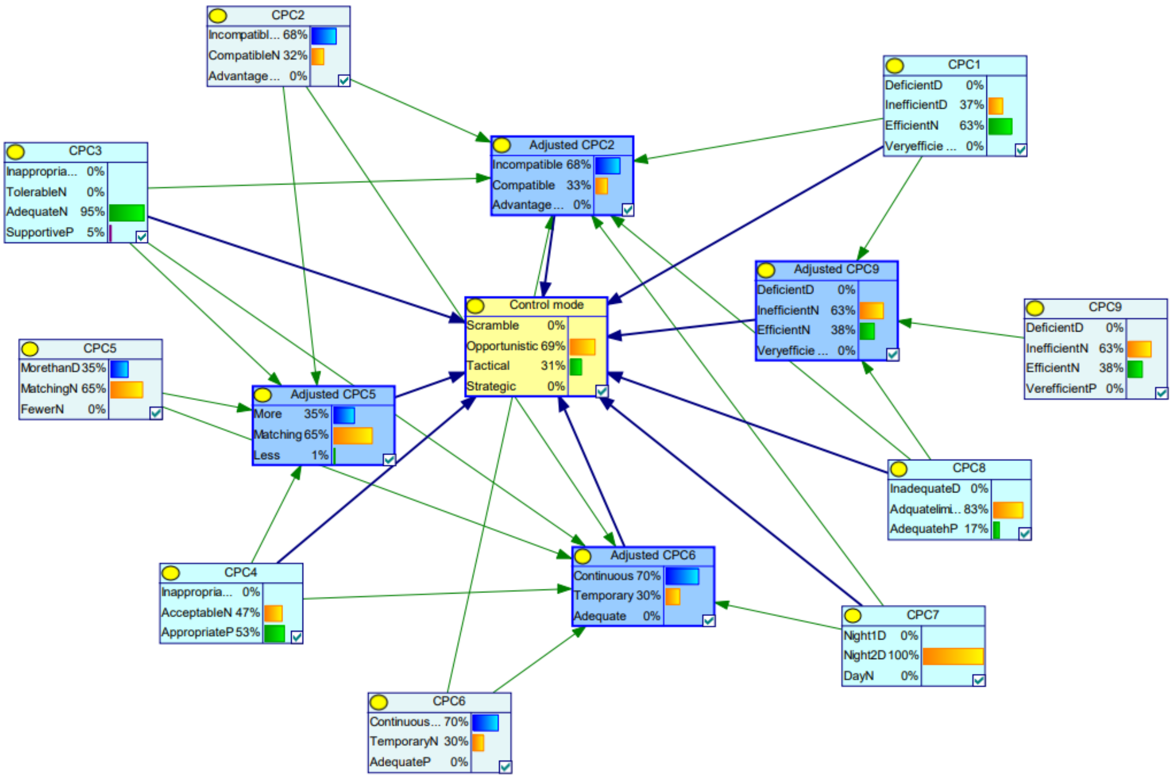

Figure 6. 14 BN-CREAM by multiple experts' judgement

6.4 Finding and discussion

The contextual factors were evaluated to predict the human error mode of emergency steering by dividing them into a single expert and expert group. As shown in ***Figure 6. 15***, the probability of human error of 3.11E-02 and 3.85E-02 and both COCOMs are located between Opportunistic and Tactical modes. The results obtained through the analysis can be used as follows. First, it identifies the vessel's current contextual mode and provides a rough assessment of the vessel's failure likelihood for the planned operations. This is the outcome of thoroughly examining the ship's current resources, individual, organisational, and environmental factors. It can assist in determining whether an emergency training or operation is appropriate for the current situation the ship encounters. Second, it can be used as a screening tool to see if further risk analysis including human reliability assessment, is required. As a result, the outcome of the analysis for the case study might be interpreted as indicating that an emergency steering drill can now be conducted, although more risk analysis is required to strengthen the ship's safety system. Interestingly, these results showed no noticeable difference between single expert judgement and multiple expert judgement. However, there is a distinct possibility that a single expert's assessment will show significant deviations due to the expert's selection and bias. As previously stated, a single expert's approach prepares for unavoidable scenarios that require significant and immediate results. Despite its obvious limitations, it is intended to provide analysts with a variety of possibilities. Therefore, further caution must be exercised when employing a single expert.

The proposed method provides a simple tool to quickly estimate the overall risk level of a ship when resources for risk analysis are very limited, such as the captain's sole decision. However, this method has clear limitations that do not provide results on which tasks are more dangerous or how much each PSF contributes because the proposed method is designed to be used as quickly and efficiently as possible by simplifying the task analysis process and weighting process for PSFs.

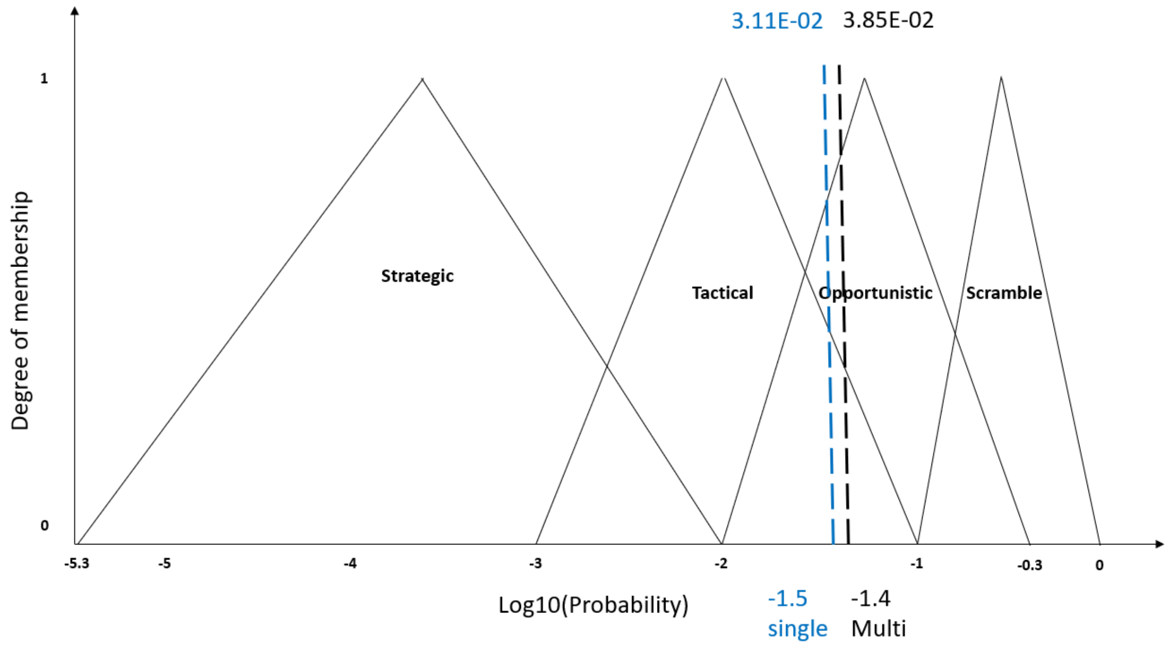

Figure 6. 15 COCOM and Human error probabilities per single and multiple experts

6.5 Chapter summary

In this chapter, the method of estimating the human error mode using BN-CREAM is described by dividing it into a single expert and expert group for the emergency steering operation. This chapter is presented to help understand the comprehensive analysis of human reliability assessment, which will be further described in Chapters 7 to 9 and will deal with full-scale human reliability assessment with a more specific task analysis.

7 Human error quantification for human-oriented tasks

7.1 Chapter overview

The primary question for a human reliability assessment in the maritime domain is, how we can accurately comprehend the human factors present in a maritime operation on a practical level. This chapter discusses a novel approach that is based on Cognitive Reliability and Error Analysis (CREAM). The method's defining feature is that it establishes a framework for evaluating specific scenarios involving maritime human errors and for assessing the context in which human actions emerge. The output of the context assessment is then to be used as model inputs for reflection of the context effect in the procedure assessment. The proposed approach is divided into two components: context assessment processing and human error quantification modelling. To increase the reliability of human error quantification, this study adopted a Fuzzy multiple attribute group decision-making method, Bayesian networks, and evidential reasoning. The Fuzzy conclusion of the context assessment is used as the model input in the CREAM basic method and as weighting factors in the CREAM extended method to account for the fact that the probability of human failure varies with external conditions. The chapter is organised in the following manner to accomplish this goal: Section 7.2 discusses human factors in the maritime industry and gives an overview of CREAM. Section 7.3 describes the proposed method, and section 7.4 presents a case study of a ship's engine room firefighting procedures. The findings and discussion are presented in Section 7.5, followed by a chapter summary in Section 7.6.

7.2 Human factors in the maritime industry and CREAM overview

While safety is a critical issue in maritime, it remains challenging to predict and prevent accidents because the cause of the accident can be a variety of factors. Notably, the human factors associated with ship operation in the maritime industry played a significant role in the accident. Human error is strongly associated with accidents, accounting for between 65 and

90% of all accidents (Kristiansen (2013); Ung (2015); Akyuz et al. (2018); Kurt et al. (2016b); Antão and Soares (2019)). According to Kurt et al. (2015) and Kurt et al. (2016a), their research conducted in the EU funded SEAHORSE Project concluded that 20-30% of standard operating procedures are ineffective hence not being followed strictly during operations. This means we need to bring more attention to review procedures on board with a specific focus on human performance to achieve safer operations. However, the terms '*human factors*' and '*human error*' are often used without a clear understanding (Khan, 2008). It is since the seafarers face many hazardous situations since they should not only carry out the navigation of ship but also have to conduct other responsibilities such as cargo loading and discharging, ballasting and de-ballasting, bunkering and maintenance work including hot and closed space work mostly independently in space away from land. Specific parts of the ship's functions have been automated, but a human still controls or interacts with most of the work. Therefore, to ensure safety at sea, human factors, specifically Human Reliability Analysis (HRA), needs to be considered at the core of safety assessments. However, HRA has always been a concern for safety engineers and risk assessment analysts due to the fundamental limitations such as insufficient data, methodological limitations related to subjectivity of analysts and expert judgment, and uncertainty concerning the actual behaviour of people during accident conditions (Konstandinidou et al., 2006). According to Schröder-Hinrichs et al. (2011), it is more difficult to collect reliable data because human and organisational factors related to accident development and emergency response are not reported enough. In this context, prospective methods for quantifying human reliability across the first generation and over the third generation HRA methods have been proposed through the nuclear and aviation sectors and recently applied to the maritime sector, but the third generation methods are still in the development stage. As a representative method, cognitive reliability and error analysis method (CREAM) was first developed by Hollnagel (1998) and can be considered as one of the most popular and commonly used second generation HRA method.

According to studies conducted by Hollnagel (1998) and later by Fujita and Hollnagel (2004), to predict human performance reliability, a context description must be provided because a discussion of what is likely to happen in a given situation must be based on a description of the specific circumstances or conditions. Therefore, it is reasonable that human error probability can be determined directly from a characterisation of the context. This condition is described in terms of the degree of control presented by four characteristic control modes: Strategic, Tactical, Opportunistic and Scrambled modes, which identify different performance reliability. The CREAM can be used for both retrospective and prospective purposes, and CREAM can apply to qualitative and quantitative analysis. The quantitative CREAM consists of basic and extended methods. Firstly, the CREAM basic method is a human failure probability quantification process that defines nine conditions, such as working conditions and crew collaborations, called Common Performance Conditions (CPCs) affecting human performance. In a basic predictive CREAM, it evaluates CPCs to predict human error probability concerning the contextual control modes with four different failure probability intervals corresponding to a value of combined CPC scores by using mapping in the diagram of the control mode. This method is mainly used for screening purposes in HRA and can identify conditions that may reduce or improve the human reliability aspects of risk assessment. While subsequent and more detailed analyses of human interactions can be acquired by the CREAM extended method (He et al., 2008), the combined score of the CPCs for context assessment derived from the basic method can be an essential parameter for the extended method. Therefore, the extended method will be able to obtain more accurate results for designated tasks of the procedures.

In this regard, this chapter proposes a new framework for estimating human error probabilities through scenario description and procedure analysis based on the CREAM method and illustrates the practical application by offering a way to transform human activities on board and their contextual conditions into analytical forms for HRA.

7.3 Methodology

This section introduces a hybrid approach combining Fuzzy theory, Bayesian network and evidential reasoning to CREAM to predict human error probability in maritime onboard procedures. The Fuzzy multiple attributive group decision-making methodology by Ölçer and Odabaşi (2005) is employed and customised for opinion aggregation to minimise the subjectivity of experts' judgment. According to Marseguerra et al. (2007), human performance in accidents has shown that the influence of the contextual conditions on the task is greater than the characteristics of the task itself. The context of a critical maritime scenario which may include factors such as time management, the external environment, proper procedures, and the crew training level, is more important and safety-critical in an emergency when compared to typical operating situations. Therefore, the effect of the context should be considered when predicting human error. In this respect, the CREAM method is selected as an appropriate framework for evaluating maritime emergency procedures on ships. The reasons are that firstly, CREAM can be used to assess the context and apply it to an analysis of cognitive activities required for individual tasks, respectively. Secondly, CREAM is a convenient structure to employ other techniques for developing an advanced approach. The flow chart of the process is shown in ***Figure 7. 1***.

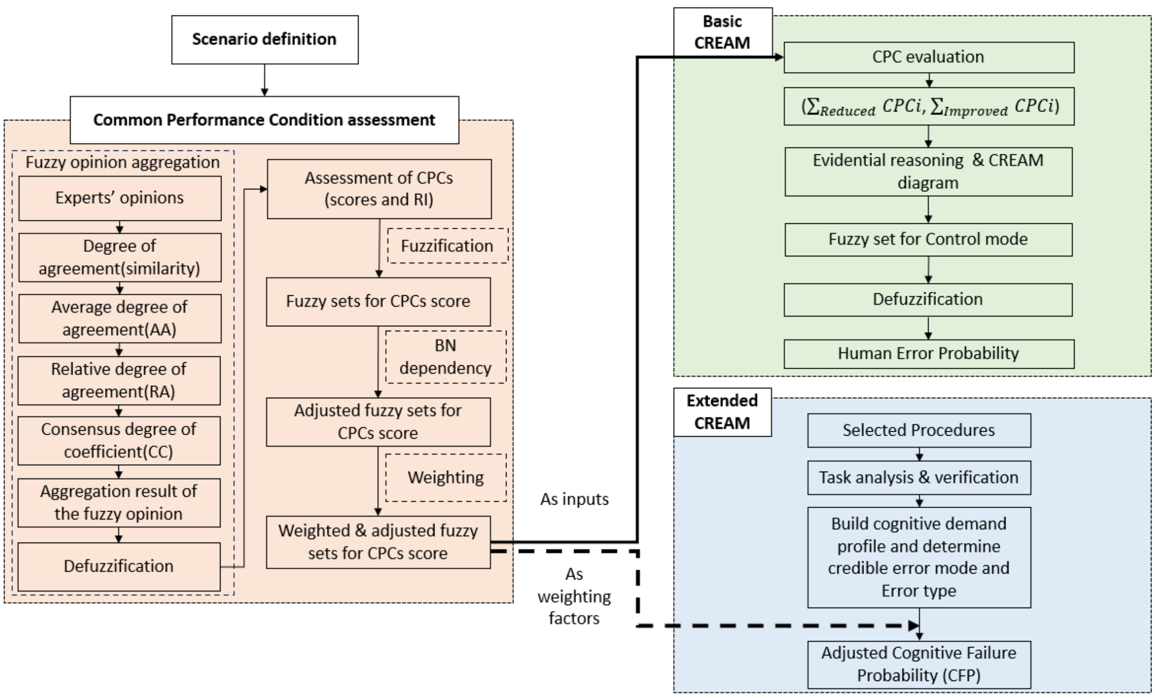

Figure 7. 1 Flow chart of the proposed approach

7.3.1 Common performance condition assessment

Individual CPCs have linguistic variables which indicate the level of CPC that addresses an expected effect on performance reliability in terms of negative or positive aspects. In the original CREAM, the only linguistic variable is decided with a 100% degree of belief for an assessment of the concerned CPC. However, a limited number of linguistic variables is insufficient to reflect CPC's impact on human reliabilities in a practical situation. To better depict the effects of CPC, Fuzzy sets are employed because Fuzzy sets are one of the best practices to tackle the ambiguity and vagueness of human error detection problems (Akyuz, 2016). Each CPC associates three or more Fuzzy sets to describe the impact of each CPCs. The Fuzzy trapezoidal number is adopted, and the corresponding Fuzzy numbers to each CPC level are developed and illustrated in **Table 6.** ***2*** of Chapter 6. The Fuzzy trapezoidal number is selected since it is intuitively easy to be used by decision-makers (Ölçer and Odabaşi, 2005). For example, 'Adequacy of organisation' is assessed with four linguistic variables, namely 'Deficient', 'Inefficient', 'Efficient' and 'Very Efficient'. The horizontal axis represents a numerical score of this CPC that varies from 0 to 100, where the most negative value is 0, and the positive is 100. The vertical axis represents a degree of membership from 0 to 1 in ***Figure 7. 2***. Note that the Fuzzy set for each CPC in this study is not an absolute value; it varies depending on the various situations and expert opinions.

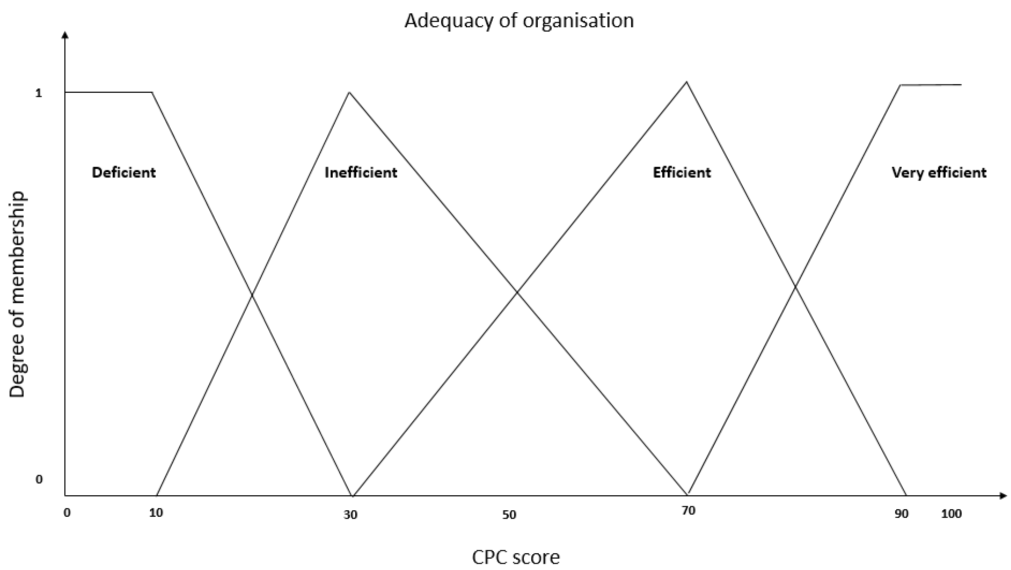

Figure 7. 2 Membership functions for Adequacy of organisation

7.3.1.1 Experts' judgement and Fuzzy opinion aggregation

The experts must assess each CPC score and their relative importance with corresponding linguistic terms. Linguistic scale for CPC level and their corresponding Fuzzy set developed and provided in **Table 6. *2*** of Chapter 6. For the relative importance of CPCs, scale and standardised Fuzzy sets are listed in ***Table 7. 1***.

Table 7. 1 Linguistic terms and their standardised Fuzzy set

Linguistic terms	Standardised Fuzzy sets
Very highly important	(0.8, 0.9, 1, 1)
Highly important	(0.6, 0.75, 0.75, 0.9)
Important	(0.3, 0.5, 0.5, 0.7)
Less important	(0.1, 0.25, 0.25, 0.4)
Not related	(0, 0, 0.1, 0.2)

The purpose of the application of the Fuzzy opinion aggregation is to translate the experts' multiple qualitative assessments of CPC score and relative importance into a single aggregated opinion with Fuzzy opinion and convert it into a crisp value through defuzzification. The opinion aggregation procedure is made based on a Fuzzy multiple attributive group decision-making methodology by Ölçer and Odabaşi (2005) and modified as follows;

(a) Calculate the degree of agreement (Similarity)

Let's assume that A=(a_1, a_2, a_3, a_4), B=(b_1, b_2, b_3, b_4) and A and B are standardised Fuzzy sets. Here, S(A, B), which is the degree of similarity between A and B, is measured by the below equation;

$$S(A,B) = 1 - \frac{|a_1 - b_1| + |a_2 - b2_1| + |a_3 - b_3| + |a_4 - b_4|}{4} \tag{7-1}$$

(b) Calculate the average degree of agreement (AA)

Let's define AA(Ex_i) as the i-th average degree of agreement and calculated by equation 2 as bellows;

$$AA(Ex_i) = \frac{1}{D-1} \sum_{\substack{i=1 \\ i \neq j}}^{D} S(Ex_i, Ex_j) \tag{7-2}$$

Where D is a number of experts

(c) Calculate the relative degree of agreement (RA)

Let's define RA(Ex_i) as the i-th relative degree of agreement and calculated by equation 3 as bellows;

$$RA(Ex_i) = \frac{AA(Ex_i)}{\sum_{i=1}^{D} AA(Ex_i)} \tag{7-3}$$

(d) Calculate the consensus degree coefficient (CC)

Let's define CC(Ex_i) as the consensus degree coefficient for i-th expert and calculated by equation 4 as bellows;

$$CC(Ex_i) = \beta * w_i + (1-\beta) * RA(Ex_i) \tag{7-4}$$

Where β is a relaxation factor between 0 and 1. A Homogeneous group of the expert is considered when β is 0 (Ölçer and Odabaşi, 2005). A coefficient w_i means the relative importance among the different experts.

(e) Calculate the aggregation result of the Fuzzy opinion (R_{AG})

The aggregated result of the experts' judgement R_{AG} can be obtained as

$$R_{AG} = \sum_{i=1}^{D} CC(Ex_i) * P(Ex_i) = (S_1, S_2, S_3, S_4) \tag{7-5}$$

(f) Defuzzification

Finally, Fuzzy opinions (R_{AG}) for each CPC and their relative importance are converted to crisp value by a centre of gravity (COG) method (Takagi and Sugeno, 1985) as

$$x = \frac{\int_{S_1}^{S_4} \mu(x) * x \, dx}{\int_{S_1}^{S_4} \mu(x) \, dx} \tag{7-6}$$

Noted that defuzzified CPC scores need to be converted from standardised numbers to their original score with an interval between 0 and 100 and the relative importance of CPC (RI_i) is a normalised number that means $\sum_{i=1}^{9} RI_i = 1$.

7.3.1.2 Fuzzification

Based on the defuzzified aggregated experts' opinion for the CPC level, the CPC scores are associated with a Fuzzy set to the CPC level.

Let L_{ij}, μ_{ij} and CPC_i define as follows.

L_{ij} represents a j-th linguistic variable for i-th CPC.

μ_{ij} is a value of membership for L_{ij}.

CPC_i is a belief structure corresponding to the i-th CPC score and expressed as follows.

$CPC_i = ((\mu_{i1}, L_{i1}), (\mu_{i2}, L_{i2}), (\mu_{i3}, L_{i3}), (\mu_{ij}, L_{ij}))$, where i= [1, 9] and j=[1, 4] (7-7)

The Fuzzy trapezoidal set expressed as (a,b,c,d) and membership function μ_{ij} for random score x is obtained by equation (6-1) of Chapter 6.

7.3.1.3. Adjusted belief structure for CPC

In the previous step, each CPC is expressed by a belief structure. However, the relation of dependency among CPCs should be considered, and CPCs are to be adjusted because CPCs are not independent of the effect of other CPC. The rules for the mutual impact of CPCs are defined as shown in ***Table 6. 4*** of Chapter 6. For example, the Rule of the 4th row indicates that 'Crew collaboration quality' depends on both 'adequacy of organisation' and 'adequacy of training and experience'. If 'crew collaboration of quality' is inefficient (Neutral) AND 'Adequacy of organisation' is very efficient (Positive) AND 'Adequacy of training and experience' is Adequate, high experience (Positive), then "Crew collaboration quality is adjusted to positive from neutral. Interactive relations can be modelled by a Bayesian network technique (Yang et al., 2013) and enable presenting rather complex systems (Hänninen, 2014). Bayesian network model based on Rules acquires four new adjusted CPCs from the nine original CPCs. Adjusted CPCs are also represented by a new belief structure as follows.

$CPC_i' = ((\mu_{i1}', L_{i1}), (\mu_{i2}', L_{i2}), (\mu_{i3}', L_{i3}), (\mu_{ij}', L_{ij}))$, where i= [1, 9] and j= [1, 4] (7-8)

Nine CPCs enter into a model as input variables with belief structures, and 4 CPCs are adjusted based on dependency rules.

7.3.1.4 Weighted Fuzzy set of CPC_i

An important issue regarding the model is whether all input parameters have equal importance (Konstandinidou et al., 2006) because the distinction of CPCs is not assumed to be independent of one another (Fujita and Hollnagel, 2004). Therefore, the relative importance of CPCs must be considered in the assessment process and carefully decided by expert judgement. The relative importance of each CPC was assigned by expert judgment in section 7.3.1.1. So, this section explains how to apply a relative importance value from the expert judgement to the proposed framework. For calculation purposes, it is needed to define a weighting factor W_i which is calculated by multiplying the number of CPCs (i.e. 9) by RI_i. The adjusted & weighted CPCi" from the original assessment of CPC score is represented as follows by multiplying weighting factors to adjusted CPCi'.

$$W_i = 9 \times RI_i \quad (7\text{-}9)$$

$$\mu_{ij}'' = W_i \times \mu_{ij}' \quad (7\text{-}10)$$

$$CPC_i'' = ((\mu_{i1}'', L_{i1}), (\mu_{i2}'', L_{i2}), (\mu_{i3}'', L_{i3}), (\mu_{ij}'', L_{ij})), \text{ where } i= [1, 9] \text{ and } j= [1, 4] \quad (7\text{-}11)$$

7.3.2 Human error quantification with the CREAM basic method

This section describes the process of determining the significant contextual control mode and predicting overall human failure probability in the specific scenario by utilising nine Fuzzy sets resulting from the context evaluation. The method consists of three main steps. First, nine Fuzzy sets are combined with positive and negative CPC scores. These two crisp values indicate the point (sums of the reduced CPCs, sums of the improved CPCs) on the two-dimensional CREAM Diagram of Control Mode in ***Figure 7. 3***. Secondly, the control mode corresponding to the point of combined CPC score is determined with a form of the Fuzzy set for four control modes through evidential reasoning. Finally, the human error probability is obtained through a defuzzification process by the Weighted Mean of Maxima method from the Fuzzy set of control mode.

7.3.2.1 CPC evaluation

Fuzzy sets of CPCs scores can be quantified to a numerical value by defining a specific value as follows.

$$L_{ij} = \begin{cases} 1, & L_{ij} \text{ is 'Improved'.} \\ 0, & L_{ij} \text{ is 'Not significant'.} \\ -1, & L_{ij} \text{ is 'Reduced',} \end{cases} \qquad (7\text{-}12)$$

$$CPC_i'' = \sum_{j=1}^{n} \mu_{ij}'' * L_{ij}, \text{ where n= 3 or 4} \qquad (7\text{-}13)$$

CPC_i'' value has one of three values depending on the expected number: positive number, negative number, or zero. To combine the CPC score, positive numbers are added separately between positive numbers and negative numbers. For not significant cases, i.e. L_{ij}=0, it is possible to assume $\sum_{Not\ significant} CPC_i''$will not make a severe difference (Hollnagel, 1998) and does not need to be considered. The combined CPC score is finally represented on the Cartesian coordinate system in the form as ($\sum_{Reduced} CPC_i''$, $\sum_{Improved} CPC_i'$)

7.3.2.2 Fuzzification of combined CPC score

The Contextual Control Mode (COCOM) is the output for nine performance condition assessments. Human error probability concerning four control modes is defined with Fuzzy triangular sets, as shown in ***Figure 6. 11*** based on the control modes and action probability in

Table 6. 8. The human error probability is represented by the Napierian logarithm function. The combined CPC_i score is regarded as a point on the diagram of the CREAM methodology for operator control mode, as shown in ***Figure 7. 3***. However, the original diagram of CREAM provides four different control modes with their error probability interval in

Table 6. 8. For the specific human error probability estimation corresponding to each combined CPC_i scores, the approach introduced by Yang et al. (2013) based on the evidential reasoning algorism of Jian-Bo and Dong-Ling (2002) is employed to infer the distribution of degrees of belief to four control modes from a basic diagram of CREAM for operator control modes in this paper. This method avoids the problem of incorporating Fuzzy logic into CREAM because too many If-Then Rule need to be established in the inference engine(Wu et al., 2017). In the proposed method, the control mode of the selected scenario is estimated by the distribution of degrees of belief to the four control modes instead of the single control mode in a logical way. The algorithm of human error probability estimation to a point K of the combined CPC score can be analysed and explained by the following pathways. Let point K to be corresponding to the combined CPC score, ($\sum_{Reduced} CPC_i{}''$,$\sum_{Improved} CPC_i{}'$), defined as the coordinates of x and y on the diagram, as shown in ***Figure 7. 3***. The distribution of degrees of belief corresponding to four control modes consisting of Strategic (D_1), Tactical (D_2), Opportunistic (D_3) and Scrambled (D_4) is defined by a set A^K and represented as follows.

$A^K = ((A^k_1, D_1), (A^k_2, D_2), (A^k_3, D_3), (A^k_4, D_4))$, where $\sum_{i=1}^{4} A_i^k = 1$ (7-14)

The set of A^K can be obtained by synthesising two different subsets of the distribution of control mode, A^{K-} and A^{K+}, which are obtained by analysing the portion of squares of varying control modes in each row and a column about the point K as shown in ***Figure 7. 3*** and expressed by as follows.

$A^{K-}= ((A^{k-}_1, D_1), (A^{k-}_2, D_2), (A^{k-}_3, D_3), (A^{k-}_4, D_4))$

$A^{K+}= ((A^{k+}_1, D_1), (A^{K+}_2, D_2), (A^{k+}_3, D_3), (A^{k+}_4, D_4))$

Where $\sum_{i=1}^{4} A_i^{K+} = 1,\ \sum_{i=1}^{4} A_i^{k-} = 1$ (7-15)

The difference between synthesising process introduced by Yang et al. (2013) and the proposed method is not to define the whole If-Then Rule, but to represent the selected CPC score into a distribution of belief degrees to the four control modes for quantification by defuzzification. The process to derive set A^K from A^+ and A^- is as follows.

Firstly, suppose coefficient values, θ^{K+} and θ^{K-}, represent a normalised number as equation (17) corresponding to X = ($\sum_{Reduced} CPC_i$ ´´+1) and Y = ($\sum_{Improved} CPC_i$ ´´ +1) from point K. The reason for adding one respectively to the sum of positive and negative CPC is that the centre of the coordinates is moved parallel from (0,0) to (1,1) to prevent the normalised value θ from being zero when both $\sum_{Reduced} CPC_i$ ´´ and $\sum_{Improved} CPC_i$ ´´ are zero on the diagram.

$\theta^{K-} = \frac{X}{X+Y}$, $\theta^{K+} = \frac{Y}{X+Y}$ (7-16)

Then, assume that M^{K+} and M^{K-} are sets of belief degrees to support the hypothesis that the set A^{K+} and A^{K-} are identified in four control modes. It means a higher score of improved CPC increases the value of θ^{K+} and a higher score of reduced CPC increases the value of θ^{K-}, thus sets M^{K+} and M^{K-} support the hypothesis of set A^{K+} and A^{K-} respectively as weights.

M^{K-} = (($\theta^{K-}A^{k-}{}_1$, D_1), ($\theta^{K-}A^{K-}{}_2$, D_2), ($\theta^{K-}A^{k+}{}_3$, D_3), ($\theta^{K-}A^{k-}{}_4$, D_4))

M^{K+} = (($\theta^{K+}A^{k+}{}_1$, D_1), ($\theta^{K+}A^{K+}{}_2$, D_2), ($\theta^{K+}A^{k+}{}_3$, D_3), ($\theta^{K+}A^{k+}{}_4$ D_4)) (7-17)

Finally, an output of the human error quantification model is represented as a set A^K = ($A^k{}_1D_1$, $A^k{}_2D_2$, $A^k{}_3D_3$, $A^k{}_4D_4$). It is a distribution of belief degrees to the four control modes for four control modes against a random point K which have $\sum_{Reduced} CPC_i$ ´´and $\sum_{Improved} CPC_i$ ´´ in the selected scenario and relevant coefficients and equations are followed.

$A_i^{k\prime}$ = P(M_i^{K+} x M_i^{K-} + M_i^{K+} x θ^{K+} + M_i^{K-} x θ^{K-})

H = P (θ^{K+} x θ^{K-})

$P = \left|1 - \sum_{T=1}^{4} \sum_{R=1,R\neq T}^{4} (M_T^{K+} * M_R^{K-})\right|^{-1}$

$A_i^k = \frac{A_i^{k\prime}}{1-H}$, (i =1, 2, 3, 4)

$A^K = ((A^k_1, D_1), (A^k_2, D_2), (A^k_3, D_3), (A^k_4, D_4))$ (7-18)

Where H is the non-normalised remaining belief unassigned after the commitment of belief to the four control modes as a result of the synthesis of A^+ and A^- and P is the normalising factor.

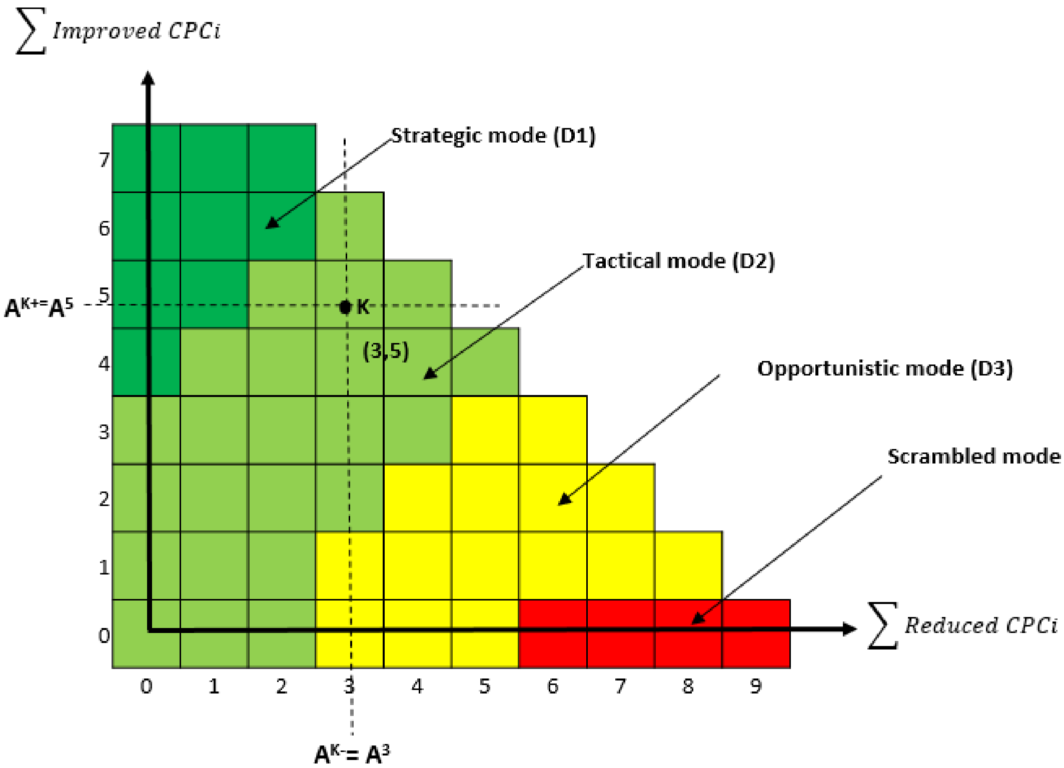

Figure 7. 3 CREAM diagram of control mode

7.3.2.3 Defuzzification and human error probability

Defuzzification is the process of converting a Fuzzy conclusion into a crisp value. The Weighted Mean of Maxima (WMoM) method is selected for this defuzzification. A set of belief degrees to the four control modes is defuzzified into a crisp value as follows.

Crisp value (CV) = $\sum_{i=1}^{4} A_i^k * w_i$ (7-19)

Where w_i is the significant value of the i-th Fuzzy membership function.

The weighted value of a Fuzzy membership function is a transverse axis when the Fuzzy membership function is a maximum value. Membership functions have been developed based on CREAM's human failure probability interval, as shown in ***Figure 6. 11***. The value w_i can be calculated as -3.651, -2, -1.151 and -0.5. The final step is to convert a crisp value to human error probability since the CV is a logarithm value of human failure probability as below.

$$HEP \text{ (human error probability)} = 10^{CV} \quad (7\text{-}20)$$

In the proposed method, all points on the surface can represent individual human error probability corresponding to the combined CPC scores, contrary to the conventional method addresses four modes for the 52 sets of CPC scores. This method makes the quantitative model more sensitive to the changes in the input value.

7.3.3 Human error quantification with the CREAM extended method

The CREAM extended method aims to produce specific action failure probabilities (Hollnagel, 1998), while the basic method does not consider specific human activities in predicting the action failure probability but only through a context assessment. The CREAM extended method can be applied if further analysis is required through the screening process using the human error probability obtained through the CREAM basic method or when the study of individual event sequences is desired. Regarding risk assessment, this method can also be utilised for procedure review by identifying the delicate tasks needing risk control options or a task to revise from the whole procedure. The CREAM extended method consists of three main steps, and the basic framework in this paper follows the original CREAM extended method introduced by Hollnagel (1998). The significant characteristic of the proposed method is that weighted and adjusted Fuzzy sets for CPC scores are utilised to adjust a nominal cognitive failure probability. Therefore, this section summarises task analysis and verification in the step. 1, building the cognitive demand profile and determining the credit failure mode in step. 2, then describes how to use Fuzzy sets to adjust the cognitive failure probabilities.

7.3.3.1 Task analysis and verification

Task analysis refers to methods of formally describing and analysing human-system interaction (Kirwan, 1994). Task analysis is conducted to define the steps which address the designated duties that the crew should complete successfully to achieve the main goal of the procedures with a hierarchical task analysis from the selected scenario. Then, the equipment or procedures of a vessel shall be evaluated to ensure that it satisfies the compulsory requirements of the domestic law or international convention according to the navigational area due to its operational characteristics. This process requires identifying the relevant provisions of the international and domestic conventions to verify the procedures' suitability.

7.3.3.2 Build a cognitive demand profile and determine a credible error mode

This step starts with describing the scenario according to the event sequence and identifying cognitive activities that characterise the activity of each work stage or event segment. The fifteen cognitive activity types are provided, and each cognitive activity is associated with one or more basic cognitive functions that consist of observation, interpretation, planning and execution by a generic cognitive-activity-by-cognitive-demand matrix as shown in ***Table 7. 2***. Once the cognitive demand is decided for the task element, the next step is to identify the most likely generic failure type for the cognitive activity of the task element. The four basic cognitive functions are classified into 13 generic failure types, and the corresponding cognitive failure probability (CFP) for each generic failure type is given, as shown in ***Table 7. 3***.

Table 7. 2 Generic cognitive activity by cognitive demand matrix (Hollnagel, 1998)

Cognitive Activity type	Observation	Interpretation	Planning	Execution
Co-ordinate			V	V
Communicate				V
Compare		V		
Diagnose		V	V	
Evaluate		V	V	
Execute				V
Identify		V		
Maintain			V	V
Monitor	V	V		
Observe	V			

Plan			V	
Record		V		V
Regulate	V			V
Scan	V			
Verify	V	V		

Table 7. 3 Nominal values and uncertainty bounds for cognitive function failures (Hollnagel, 1998)

Cognitive function	Generic failure type	Lower bound (5%)	Basic value	Upper bound (95%)
Observation	O1. Wrong object observed	3.00E-04	1.00E-03	3.00E-03
	O2. Wrong identification	2.00E-02	7.00E-02	1.70E-02
	O3. Observation not made	2.00E-02	7.00E-02	1.70E-02
Interpretation	I1. Faulty diagnosis	9.00E-02	2.00E-01	6.00E-01
	I2. Decision error	1.00E-03	1.00E-02	1.00E-01
	I3. Delayed interpretation	1.00E-03	1.00E-02	1.00E-01
Planning	P1. Priority error	1.00E-03	1.00E-02	1.00E-01
	P2. Inadequate plan	1.00E-03	1.00E-02	1.00E-01
Execution	E1. Action of wrong type	1.00E-03	3.00E-03	9.00E-03
	E2. Action at wrong time	1.00E-03	3.00E-03	9.00E-03
	E3. Action on wrong object	5.00E-05	5.00E-04	5.00E-03
	E4. Action out of sequence	1.00E-03	3.00E-03	9.00E-03
	E5. Missed action	2.50E-02	3.00E-03	4.00E-02

7.3.3.3 Adjusted CFP by weighting factors

The last step in the CREAM extended method is to adjust the nominal CFP concerning the effect of the CPC. Nine Fuzzy sets for all CPC scores are utilised in this step. For example, the Fuzzy set ((μ_{11}'', L_{11}), (μ_{12}'', $L_{12,}$), (μ_{13}'', $L_{13,}$), (μ_{14}'', L_{14})) represent a Fuzzy score of CPC_1. First, let's define W_{ijn} as a weighting factor for the n-th generic failure type of the j-th CPC level at the i-th CPC and get data from the original CREAM by Hollnagel (1998). Then, let's define W_{in} as a weighting factor for n-th cognitive function of CPC_i. The weighting factor, W_n, is acquired as follows.

$$W_{in} = \sum_{j=1}^{4} \mu_{ij} * W_{ijn} \quad (7\text{-}21)$$

$$W_n = \prod_{i=1}^{9} W_{in} \quad (7\text{-}22)$$

Where i= 1 to 9, j=1 to 3 or 4 and n= observation, Interpretation, planning and Execution.

7.4 Application of Fuzzy CREAM to engine room fire drill on a ship (Case study2)

According to Darbra and Casal (2004), accidents associated with fire and explosion at seaports account for 29% and 17%, respectively. The statistical analysis for Maritime Accident Investigation Branch (MAIB) data by de Maya et al. (2019a) found that fire and explosion accidents account for 6.78% of all maritime accidents from 1990 to 2016. Those incidents have a reputation for high mortality. Weng and Yang (2015) show that fire and explosion related incidents are 132% higher in death tolls than other types of accidents. In particular, for passenger ships, fire/explosive accidents are the most frequent occurrence of a total loss of vessels compared to other accidents (Eliopoulou et al., 2016). According to Baalisampang et al. (2018), 48% of fire incidents in ships are related to human error, followed by mechanical failure at 22% and temperature response at 14%. In this context, this paper was motivated to apply the proposed method for potential fire incidents in an engine room where most fire incidents occur.

For an illustration of the proposed approach, both scenario and procedures for the engine room fire-fighting in general cargo ship have been selected since fire drill at sea is a critical situation in which the crews are required to complete tasks for fire-fighting with limited resources such as personnel, equipment and time. The scenario of an engine room fire-fighting is described in section 7.4.1 to assess CPCs and predict overall HEP without considering specific human activity in the selected control mode by the CREAM basic method. The CREAM extended method used the engine room fire drill procedure in section 7.4.4.1 to conduct task analysis and predict individual CFP to all tasks.

The application of the proposed method to case study and data collection was conducted in the following ways.

Firstly, to develop an actual emergency response procedure, the existing fire-fighting procedures used in cargo ships were obtained from numerous companies. A group of experts verified and enhanced the procedure to ensure compliance with SOLAS and STCW requirements. Next, the scenario was generated to reflect the nine CPC characteristics through meetings of the expert group. Also, a criterion was applied when selecting experts for the evaluation stage. In other words, experts with practical experience in fire-fighting drills on ships as crew members or safety system auditors are chosen for this evaluation. Then, the assessment was conducted independently of each expert to eliminate the group thinking bias. The fire-fighting procedures and scenarios were provided and evaluated by a questionnaire using linguistic terms on the relative importance of each CPC and CPC level.

7.4.1 Scenario definition

The scenario for an engine room fire drill on a general cargo ship is described to illustrate the proposed method and focuses on presenting CPCs for evaluation as follows.

On a hot summer day, a general cargo ship was waiting to depart at the anchoring position after finishing cargo loading. The temperature was 38 °C, and the humidity was 70 %. The sea conditions and winds were generally good. The vessel was five years old general cargo ship, G/T 5,000, and overall, the vessel was in good condition. The ship's management company has managed a total of 30 vessels, holding both the company's DOC certificate and SMC certificates for individual ships in accordance with an International Safety Management Code (ISM), and obtained ISO certificates on the quality management system. Last month, the company conducted an internal audit of the vessel, and all three identified nonconformities have been rectified. A total of 20 crew members were on board and were made up of three different Nationalities. Six crew members were replaced the previous day and conducted familiarisation training during the last day's afternoon. The ship’s captain planned to conduct a fire drill and abandon ship exercise on the day at 2 p.m. The fire extinguishing equipment consisted of a fixed CO2 gas system in the engine room, two main fire pumps inside the main engine room, an emergency fire pump in the steering gear room, portable fire extinguishers, and two firefighters's outfits, etc. All fire pumps were manually operated on-site and remotely in the fire control room and bridge. All fire extinguishing equipment of the ship has completed the periodical inspection in accordance with the SOLAS Convention. For communication during training, there were three portable communication devices. The company provided the Muster List to the vessel that consists of duties and responsibilities in case of such mishaps, designated and assigned to each person on the ship in case of emergency including fire and abandon ship. The captain had carried out a monthly fire-fighting and abandon ship drill three days ago, and the records were written in the ship's logbook. This drill is the first to be trained in the vessel for six newly onboard crews, while the other 14 crews joined last month's training following the captain's training plan.

7.4.2 Common performance condition assessment

Experts' relative importance is considered a heterogeneous group depending on their background and assigned as 0.20, 0.18, 0.21, 0.20 and 0.21. For assessment, experts are asked to assign CPC scores and their relative importance in ***Table 7. 4*** and ***Table 7. 5***. Then, opinion aggregation from CPC_1 to CPC_9 except for the CPC_7 and relative importance for nine CPCs are done. A relaxation factor β is assumed to be 0.5. As an example, specific aggregation for CPC4 is illustrated in ***Table 7. 6***. Finally, aggregated Fuzzy opinions are defuzzified and listed in ***Table 7. 7***. Once experts' judgement and Fuzzy opinion aggregation are completed, the next step is to convert the defuzzified CPC scores to Fuzzy membership again for a human error quantification. Then adjust Fuzzy sets by dependency relation shown in ***Figure 7. 4***, illustrated by a Genie software. Finally, the weighted & adjusted Fuzzy sets are obtained by multiplying the weighting factor by the adjusted Fuzzy sets. The Fuzzy memberships are provided in ***Table 7. 7***.

Table 7. 4 Experts' evaluations of CPCs and their standardised Fuzzy set

CPC	E1	E2	E3	E4	E5
CPC1	Efficient (0.3, 0.7, 0.7, 0.9)	Efficient (0.3, 0.7, 0.7, 0.9)	Efficient (0.3, 0.7, 0.7, 0.9)	Inefficient (0.1, 0.3, 0.3, 0.7)	Efficient (0.3, 0.7, 0.7, 0.9)
CPC2	Incompatible (0, 0, 0.2, 0.6)	Incompatible (0, 0, 0.2, 0.6)	Compatible (0.2, 0.6, 0.6, 0.9)	Compatible (0.2, 0.6, 0.6, 0.9)	Incompatible (0, 0, 0.2, 0.6)
CPC3	Adequate (0.3, 0.7, 0.7, 0.9)	Adequate (0.3, 0.7, 0.7, 0.9)	Tolerable (0.1, 0.3, 0.3, 0.7)	Adequate (0.3, 0.7, 0.7, 0.9)	Tolerable (0.1, 0.3, 0.3, 0.7)
CPC4	Appropriate (0.6, 0.9, 1, 1)	Appropriate (0.6, 0.9, 1, 1)	Acceptable (0.2, 0.6, 0.6, 0.9)	Appropriate (0.6, 0.9, 1, 1)	Acceptable (0.2, 0.6, 0.6, 0.9)
CPC5	Matching current capacity (0.2, 0.6, 0.6, 0.9)	Matching current capacity (0.2, 0.6, 0.6, 0.9)	Matching current capacity (0.2, 0.6, 0.6, 0.9)	Matching current capacity (0.2, 0.6, 0.6, 0.9)	Matching current capacity (0.2, 0.6, 0.6, 0.9)

CPC6	Temporarily inadequate (0.2, 0.6, 0.6, 0.9)	Temporarily inadequate (0.2, 0.6, 0.6, 0.9)	Temporarily inadequate (0.2, 0.6, 0.6, 0.9)	Temporarily inadequate (0.2, 0.6, 0.6, 0.9)	Temporarily inadequate (0.2, 0.6, 0.6, 0.9)
CPC8	Adequate, limited experience (0.2, 0.6, 0.6, 0.9)	Adequate, limited experience (0.2, 0.6, 0.6, 0.9)	Adequate, limited experience (0.2, 0.6, 0.6, 0.9)	Adequate, limited experience (0.2, 0.6, 0.6, 0.9)	Adequate, limited experience (0.2, 0.6, 0.6, 0.9)
CPC9	Inefficient (0.1, 0.3, 0.3, 0.7)	Efficient (0.3, 0.7, 0.7, 0.9)	Efficient (0.3, 0.7, 0.7, 0.9)	Inefficient (0.1, 0.3, 0.3, 0.7)	Efficient (0.3, 0.7, 0.7, 0.9)

Table 7. 5 Experts' evaluation of the relative importance of CPCs

	E1	E2	E3	E4	E5
RI1	Moderate (0.3,0.5,0.5,0.7)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Moderate (0.3,0.5,0.5,0.7)	Highly important (0.6,0.75,0.75,0.9)
RI2	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)
RI3	Highly important (0.6,0.75,0.75,0.9)	Moderate (0.3,0.5,0.5,0.7)	Highly important (0.6,0.75,0.75,0.9)	Moderate (0.3,0.5,0.5,0.7)	Moderate (0.3,0.5,0.5,0.7)
RI4	Moderate (0.3,0.5,0.5,0.7)	Moderate (0.3,0.5,0.5,0.7)	Very highly important (0.8,0.9,1,1)	Moderate (0.3,0.5,0.5,0.7)	Moderate (0.3,0.5,0.5,0.7)
RI5	Moderate (0.3,0.5,0.5,0.7)	Highly important (0.6,0.75,0.75,0.9)	Moderate (0.3,0.5,0.5,0.7)	Less important (0.1, 0.25, 0.25, 0.4)	Moderate (0.3,0.5,0.5,0.7)
RI6	Moderate (0.3,0.5,0.5,0.7)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Highly important (0.6,0.75,0.75,0.9)	Moderate (0.3,0.5,0.5,0.7)
RI7	Less important	Highly important	Moderate	Less important	Moderate

	(0.1,0.25,0.25, 0.4)	(0.6,0.75,0.75, 0.9)	(0.3,0.5,0.5,0. 7)	(0.1,0.25,0.25, 0.4)	(0.3,0.5,0.5,0. 7)
RI8	Highly important (0.6,0.75,0.75, 0.9)	Very highly important (0.8,0.9,1,1)	Very highly important (0.8,0.9,1,1)	Very highly important (0.8,0.9,1,1)	Highly important (0.6,0.75,0.75, 0.9)
RI9	Highly important (0.6,0.75,0.75, 0.9)	Highly important (0.6,0.75,0.75, 0.9)	Highly important (0.6,0.75,0.75, 0.9)	Moderate (0.3,0.5,0.5,0. 7)	Less important (0.1,0.25,0.25, 0.4)

Table 7. 6 Aggregation under the CPC$_4$

Fuzzy sets for expert opinion					**Relative degree of agreement (RA)**	
Ex1	0.60	0.90	1.00	1.00	RA(Ex1)	0.21
Ex2	0.60	0.90	1.00	1.00	RA(Ex2)	0.21
Ex3	0.20	0.60	0.60	0.90	RA(Ex3)	0.19
Ex4	0.60	0.90	1.00	1.00	RA(Ex4)	0.21
Ex5	0.20	0.60	0.60	0.90	RA(Ex5)	0.19
Degree of agreement(S)					**Relative degree of agreement (RA)**	
S12	1.00		S34	0.70	RA(Ex1)	0.21
S23	0.70		S15	0.70	RA(Ex2)	0.21
S13	0.70		S25	0.70	RA(Ex3)	0.19
S14	1.00		S35	1.00	RA(Ex4)	0.21
S24	1.00		S45	0.70	RA(Ex5)	0.19
Average degree of agreement (AA)					**Consensus degree coefficient (CC)**	
AA(Ex1)				0.85	CC(Ex1)	0.20
AA(Ex2)				0.85	CC(Ex2)	0.20
AA(Ex3)				0.78	CC(Ex3)	0.20
AA(Ex4)				0.85	CC(Ex4)	0.20
AA(Ex5)				0.78	CC(Ex5)	0.20
Rag (HT)	0.44	0.78	0.76	0.96		
Defuzzification				0.72		
Normalised score				**72.29**		

Table 7. 7 Fuzzy sets for the CPCs assessment for fire-fighting scenario

CPC evaluation				Fuzzy sets for CPC score			
CPC_i	score	RI_i	W_i	μ_{i1}	μ_{i2}	μ_{i3}	μ_{i4}
CPC_1	57	0.12	1.05	0	0.33	0.68	0
CPC_2	35	0.13	1.2	0.63	0.38	0	-
CPC_3	51	0.11	0.95	0	0.48	0.53	0
CPC_4	72	0.1	0.89	0	0.6	0.4	-
CPC_5	55	0.09	0.8	0.13	0.88	0	-
CPC_6	55	0.12	1.05	0.13	0.88	0	-
CPC_7	14h	0.08	0.7	0	0	1	-
CPC_8	55	0.15	1.36	0.13	0.88	0	-
CPC_9	52	0.11	1	0	0.45	0.55	0
Total	**-**	**1**	**9**	**9**			
Adjusted Fuzzy sets for CPC score				**Weighted & adjusted Fuzzy sets for CPC score**			
μ_{i1}'	μ_{i2}'	μ_{i3}'	μ_{i4}'	μ_{i1}''	μ_{i2}''	μ_{i3}''	μ_{i4}''
0	0.33	0.68	0	0	0.34	0.71	0
0.63	0.38	0	-	0.75	0.45	0	-
0	0.48	0.53	0	0	0.45	0.5	0
0	0.6	0.4	-	0	0.53	0.36	-
0.13	0.88	0	-	0.1	0.7	0	-
0.15	0.85	0	-	0.16	0.89	0	-
0	0	1	-	0	0	0.7	-
0.13	0.88	0	-	0.17	1.19	0	-
0.02	0.45	0.53	0	0.02	0.45	0.53	0
9				**9**			

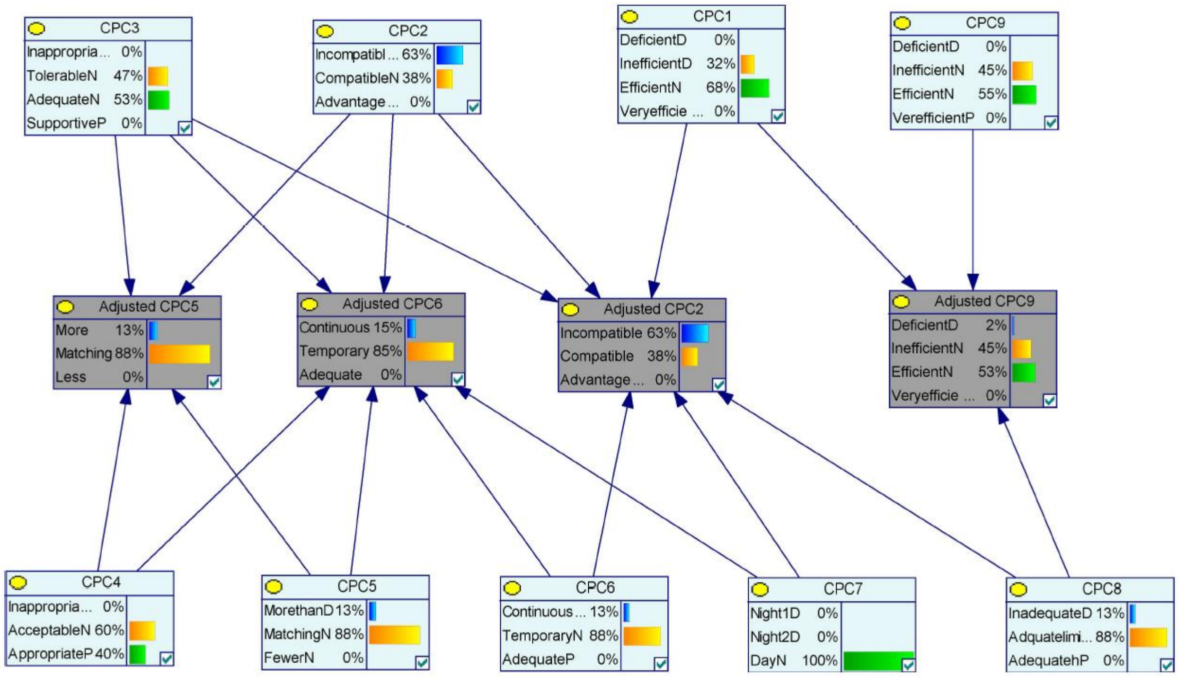

Figure 7. 4 Bayesian presentation for the dependency of the performance condition

7.4.3 Human error quantification with the CREAM basic method

This section presents the process of calculating the overall human error probability from Fuzzy memberships for CPCs using the proposed approach-based CREAM basic method.

7.4.3.1 CPC evaluation

In this step, adjusted & weighted Fuzzy sets of CPCs scores are quantified to the combined CPC score. The combined CPC score is calculated as a reduced effect of 1.54 and an improved effect of 0.36 by multiplying the expected impact following section 7.3.2.1.

7.4.3.2 Fuzzification of combined CPC score

This section describes the process of inferring the distribution of belief degrees corresponding to four control modes consisting of Strategic (D_1), Tactical (D_2), Opportunistic (D_3) and Scrambled (D_4) from the combined CPC score point K (1.54, 0.36). Subsets $A^{1.54}$ and $A^{0.36}$ are obtained by analysing the portion of squares of different control modes in each row and column to the point K as follows.

$A^{K-} = A^{1.54} = ((\frac{2}{8}, D_1), (\frac{6}{8}, D_2), (0, D_3), (0, D_4))$

$A^{K+} = A^{0.36} = ((0, D_1), (\frac{3}{10}, D_2), (\frac{3}{10}, D_3), (\frac{4}{10}, D_4))$

Normalised coefficient $\theta^{1.54}$ and $\theta^{0.36}$ are acquired after parallel movement of centre of coordinate from (0,0) to (1,1) by the equation (7-16) as follows.

$\theta^{1.54} = \frac{2.2.5418}{2.54+1.36} = 0.65$, $\theta^{0.36} = \frac{1.36}{2.54+1.36} = 0.35$

$M^{1.54}$ and $M^{0.36}$ are belief degrees to support the hypothesis that the subset A^{K-} and A^{K+} are identified in four control modes by the equation (7-17) as follows.

$M^{1.54} = ((0.65 * \frac{2}{8}, D_1), (0.65 * \frac{6}{8}, D_2), (0, D_3), (0, D_4))$

$M^{0.36} = ((0, D_1), (0.35 * \frac{3}{10}, D_2), (0.35 * \frac{3}{10}, D_3), (0.35 * \frac{4}{10}, D_4))$

Coefficients P, H and set of A^K are calculated by equation (7-18) and an output of the human error quantification model is derived as follows.

P=1.21, H=0.27

$A^{(1.54, 0.36)} = ((0.18, D1), (0.68, D2), (0.06, D3), (0.08, D4))$

7.4.3.3 Defuzzification and human error probability

A set of belief degrees to the four control modes $A^{(1.54, 0.36)}$ is defuzzified into a logarithm number negative 2.12; then HEP is derived by equation (7-20) as follows.

HEP (human error probability) = 10^{CV} = 0.0076

7.4.4 Human error quantification with the CREAM extended method

In accordance with SOLAS Chapter3, Regulation 19.3.2, all crew members shall participate in at least one abandon ship and fire drill every month (IMO, 2001). Fire-fighting facilities in each ship vary depending on the requirement of fire detection and extinguish system as well as on the type of vessels and cargo. Therefore, fire drills for specific ships should be planned so proper consideration of regular practice in various emergencies can be made. The procedures also have to consider an abandon-ship decision made by the ship's Master in case of fire-fighting failure.

7.4.4.1 Task analysis and verification

The hierarchical task analysis for the procedures of engine room fire-fighting is shown in ***Table 7. 8***. The procedures are confirmed that all compulsory requirements by SOLAS Chapter 3, Regulation 19.3.5.2 are included (IMO, 2001). The procedure consists of seven main tasks which are i) Fire detection and announcement, ii) Assembly at the muster station, iii) Check openings in the engine room area, iv) Preparation of the fireman, v) Preparation of the fire pump and water spray, vi) Fire-fighting, vii) Further actions and main tasks are divided to twenty-three subtasks as ***Table 7. 8***.

Table 7. 8 Sample procedures of the engine room fire-fighting on ships

Engine room fire-fighting procedures
1. Fire detection and announcement
1.1 Detect fire in the engine room
1.2 Report to the wheelhouse
1.3 Push the fire alarm and make an announcement
1.4 Report to stations
2. Assembly at the muster station
2.1 Ensure all crew gathered at the muster station
2.2 Check fireman's outfit and other personal rescue equipment
2.3 Describe the fire-fighting procedures and duties to all crew members
2.4 Check communication equipment
3. Check openings in the engine room area
3.1 Stop all-electric ventilation fan
3.2 Close all air inlets and doors into the engine room
3.3 Ensure no air supply into the engine room
4. Preparation of the fireman

4.1 Wear a fireman's outfit with equipment
4.2 Ensure all fireman's equipment is good in order
5. Preparation of the fire pump and water spray
5.1 Open suction valve for the fire pump
5.2 Close main isolating valve
5.3 Connect at least two fire hoses to fire hydrants
5.4 Start the (emergency) fire pump
5.5 Check the water pressure
6. Fire fighting
6.1 Start water spray to engine room boundary for cooling
6.2 Fireman, access into fire site and fire fighting
7. Further actions
7.1 Ensure fire is extinguished completely
7.2 Check the necessary of the fixed fire extinguisher system (e.g.CO2 gas)
7.3 Check the necessary of the abandon ship

7.4.4.2 Build a cognitive demand profile and determine a credible error mode

All tasks from 1.1 to 7.3 matched one of the cognitive activities associated with cognitive demand and credible failure mode. The most likely error mode for the cognitive activity of each task is decided carefully in ***Table 7. 9***. Nominal Cognitive Failure Probability (CFP_O) is provided in ***Table 7. 3***.

7.4.4.3 Adjusted CFP by weighting factors

The weighting factor per cognitive demand is calculated by equations (7-21) and (7-22) for fire-fighting procedures, and the adjusted CFP throughout the whole procedure is illustrated in ***Table 7. 9***.

Table 7. 9 CREAM extended method analysis result for the engine room fire-fighting procedures

Tasks	Cognitive activity	Cognitive Demands	The most credible Error mode	CFP_O	W_n	Adjusted CFP
1.1	Observe	Observation	O3. Observation not made	7.00E-02	2.64	1.85E-01

1.2	Communicate	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
1.3	Execute	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
1.4	Communicate	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
2.1	Observe	Observation	O3. Observation not made	7.00E-02	2.64	1.85E-01
2.2	Verify	Observation Interpretation	O2. Wrong identification	7.00E-02	2.64	1.85E-01
2.3	Communicate	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
2.4	Verify	Observation Interpretation	O3. Observation not made	7.00E-02	2.64	1.85E-01
3.1	Execute	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
3.2	Execute	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
3.3	Monitor	Observation Interpretation	O2. Wrong identification	7.00E-02	2.64	1.85E-01
4.1	Execute	Execution	E1. Action of wrong type	3.00E-03	2.98	8.94E-03
4.2	Verify	Observation Interpretation	O2. Wrong identification	7.00E-02	2.64	1.85E-01
5.1	Execute	Execution	E3. Action on wrong object	5.00E-04	2.98	1.49E-03
5.2	Execute	Execution	E3. Action on wrong object	5.00E-04	2.98	1.49E-03
5.3	Execute	Execution	E1. Action of wrong type	3.00E-03	2.98	8.94E-03

5.4	Execute	Execution	E4. Action out of sequence	3.00E-03	2.98	8.94E-03
5.5	Verify	Observation Interpretation	O2. Wrong identification	7.00E-02	2.64	1.85E-01
6.1	Execute	Execution	E5. Missed action	3.00E-03	2.98	8.94E-03
6.2	Execute	Execution	E4. Action out of sequence	3.00E-03	2.98	8.94E-03
7.1	Observe	Observation	O2. Wrong identification	7.00E-02	2.64	1.85E-01
7.2	Diagnose	Interpretation Plan	I2. Decision error	1.00E-02	3.84	3.84E-02
7.3	Diagnose	Interpretation Plan	I2. Decision error	1.00E-02	3.84	3.84E-02

7.5 Findings and discussion

The proposed approach presents individual human failure probabilities obtained by a proposed CREAM based method by separating the context assessment process and human error quantification process based on a particular maritime scenario; engine room fire-fighting procedures. From the result of the basic method, it is revealed that a significant control mode is a Tactical mode with 68 % belief and also have 18% belief in Strategic mode, 6 % belief in Opportunistic mode and 8% belief in Scrambled mode. The overall human failure probability is calculated as 0.0076, which can occur under the given circumstance described in the fire-fighting scenario. For the result of the extended method, the weighting factor per cognitive function shows the most significant adverse effect on the interpretation in a given scenario with 3.84, followed by 2.98 for execution, 2.67 for planning and 2.64 for observation. For the comparison, the weighting factor in Tactical mode is 1.90 by a simple table in the original CREAM. The range of weighting between 2.64 and 3.849 of the proposed approach is quite reasonable. The main finding is that the vulnerable subtasks with the higher failure probability are identified during the fire-fighting procedure, as shown in ***Table 7. 9***. Tasks 1.1 (Detect fire in the engine room), 2.1 (Ensure all crew gathered at the muster station), 2.2 (Check fireman's outfit and other personal rescue equipment), 2.4 (Check communication equipment), 3.3 (Ensure no air supply into the engine room), 4.2 (Ensure all fireman's equipment good in order), 5.5 (Check the water pressure), and 7.1 (Ensure fire extinguished completely) have the highest failure probability of 0.185. The lowest HEP is 0.00149 for Tasks No. 5.1 (Open suction valve for the fire pump) and 5.2 (Close main isolating valve). This result means that simple physical activity has a lower failure probability than complex cognitive activities, which need the additional ability for interpretation and decision. The study also found that 'Adequacy of training and experience' is recognised as the most significant CPC factor contributing to human error in fire-fighting scenarios with a weight of 1.36, followed by 'working conditions' with a weight of 1.20 times, 'the adequacy of organization' and ‘available time’ with a weight of 1.05. The weighting for nine CPCs is illustrated in ***Figure 7. 5***. For comparison, the original CREAM method is applied to the same assessment as ***Table 7. 10***. The overall results can be found to be within reasonable limits. Notably, the proposed method can identify the effects of other control modes that are ignored by the single control mode, and the quantified human failure probability can be obtained. The method allows the same analysis to be expressed in a more detailed output. This research result can improve the fire-fighting procedures and other critical operating procedures on the ship and contributes to safety at sea.

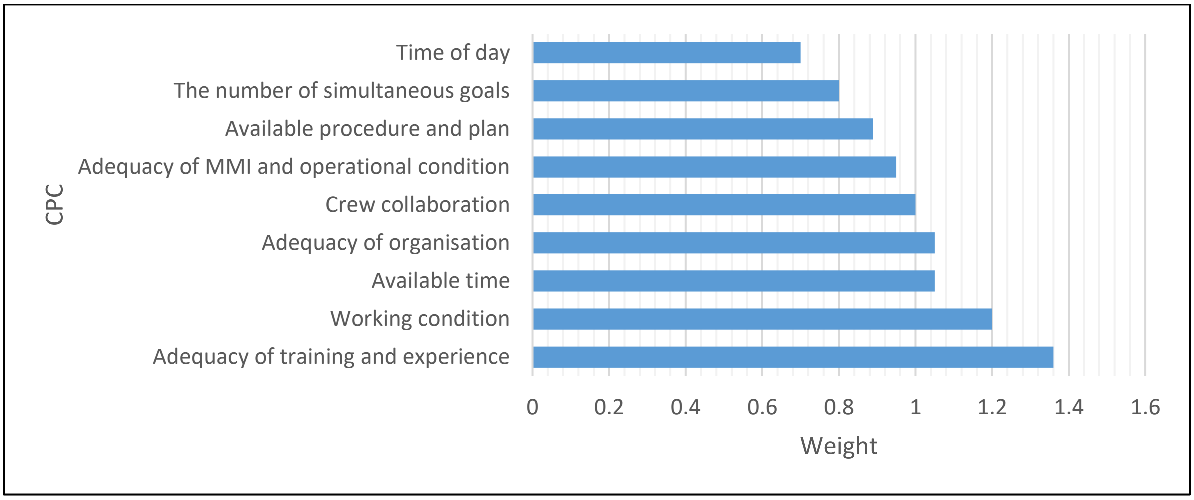

Figure 7. 5 Factors contributing to human failure in fire-fighting

Table 7. 10 Comparison result with the original CREAM method

Outputs	Original CREAM	Proposed Method
Combined CPC Score	$Expert_1$ (1,1) $Expert_2$ (1,1) $Expert_3$ (0,0) $Expert_4$ (1,1) $Expert_5$ (1,0)	(1.54,0.36)
Control Mode	Tactical (100%)	Tactical (68%) Strategic (18%) Scrambled (8 %) Opportunistic (6%)
Overall Human Error Probability	between 0.001 and 0.1	0.0076

7.6 Chapter summary

This chapter introduced a new framework-based CREAM applicable to the maritime industry and illustrated practical fire-fighting scenarios and procedures. The characteristics and expected advantages of the proposed method are: Firstly, the proposed method provides an independent process of Common Performance Condition (CPC) assessment from HEP quantification models. This structure provides a simple way to reflect a change of parameters. For example, when the concerned analysis is needed to change the type of CPCs and their linguistic terms with Fuzzy sets to reflect the characteristics of the context, the same HEP quantification model can be applied to various situations by separating the quantification model from the CPCs assessment. Furthermore, the same quantification model can be applied to individual assessments by different experts, with different weighting factors for the relative importance of CPC. This simple structure could be realised to get an instant estimation of human failure probability without adjusting the parameters of the HEP quantification model for assessing a specific task. Secondly, the output of the CPC assessment can be utilised as an input value in the CREAM basic method and weighting factors in the CREAM extended method, respectively. This method makes the whole procedure more useful by allowing the results of the CPC assessment to be used not only in the basic method but also in the extended method. Finally, the proposed method can evaluate the context in a maritime scenario based on the CREAM basic method and illustrate practical application to onboard procedures in the context of vessels using the CREAM extended method. The proposed framework also can be extended to apply to the other ship procedures with various scenarios. The quantification model does not require a rule-based inference system for a more convenient application. Instead, it infers the distribution of belief for control modes from the specific combined score of CPC for human error quantification. In conclusion, this study's results can positively impact the safety of shipping operations and the enhancement of safety at sea by providing a framework applicable to human error analysis.

8 Modelling human errors for human reliability assessment

8.1 Chapter overview

Human error mode and human errors per task were previously estimated using CREAM-based methods in chapters 6 and 7. However, these methods are not suitable for applying to relatively complex operations where machines and humans interact because CREAM-based methods do not provide a suitable model to consider interactions between errors. To expand assessment capacity from human error to human reliability, this chapter proposes an integrated model for human reliability assessment, which incorporates human errors into risk assessment. The chapter is organised in the following manner to accomplish this goal: Section 8.2 introduces research motivation and background. Section 8.3 describes the proposed method, and section 8.4 presents a case study of rescue boat drills for man overboard on ships. The findings and discussion are presented in Section 8.5, followed by a chapter summary in Section 8.6.

8.2 Research motivation and background

Once human errors are predicted, the next step is a human error representation which utilises modelling to carry out a risk assessment to reduce error. Some studies have demonstrated a risk assessment combining the human reliability assessment methods. For example, Zhou et al. (2017b) utilised the CREAM method with a modified fault tree model for LNG spill accidents during LNG carriers' handling operations for risk assessment. Ung (2019) applied fault-tree analysis where a modified Fuzzy Bayesian network-based CREAM was applied to a risk assessment of human error contribution in oil tanker collisions. Although various HRA techniques are used in maritime cases to enhance safety, the following research gaps are identified for HRA application in the maritime industry. Firstly, these HRA techniques mainly focus on quantifying human errors while they do not deal well with the dependency among tasks. Furthermore, the previous HRAs do not address how to incorporate each failure event into the system structure in a detailed method. Therefore, systematic modelling, including human error, needs to be developed. Secondly, the issue of uncertainty and inconsistency in expert judgment arising from the process of quantifying human errors remains to be improved despite efforts in previous studies. Finally, as mentioned above, SPAR-H is a technology developed for the nuclear industry, so the provided PSFs are from the operation of the nuclear industry and need to be customised for application to specific operations maritime.

Thus, a hybrid method that combines SPAR-H and a reliability block diagram is developed to fill the research gap. First, a modified SPAR-H is employed to estimate the probability of human error. Then, the Fuzzy opinion aggregation method improves research consistency while reducing subjectivity and ambiguity. Finally, the Reliability Block Diagram (RBD) analysis is used to model human reliability to consider system configuration and task dependencies. The procedures of a rescue boat drill for a man overboard in a specific context defined by a scenario are chosen to present human error probabilities for each task and human reliability for the entire procedure.

8.3 Methodology

This section introduces a hybrid approach that combines SPAR-H with the Fuzzy theory to assess human reliability during onboard procedures. To minimise the subjectivity and variability of experts, we adapt and customise Fuzzy multi-attribute group decision-making methodologies by Ölçer and Odabaşi (2005) for an opinion aggregation. The context of critical maritime scenarios may include factors like the human-machine interfaces, the complexity of the task, working conditions, and crew training levels. However, different operations are not carried out in the same environment. Therefore, the characteristics of each task and the factors that affect its performance should be evaluated individually. For this reason, the Petro-HRA method by Bye et al. (2017), based on SPAR-H by Blackman et al. (2008), is selected as an appropriate framework for evaluating ship offshore emergency procedures. This is because the SPAR-H method helps measure the effectiveness of performance shaping factors on human performance for individual tasks to estimate human errors. At the same time, the Petro-HRA provides a comprehensive quantitative risk assessment framework for whole procedures. The flowchart of the proposed approach is shown in ***Figure 8. 1***.

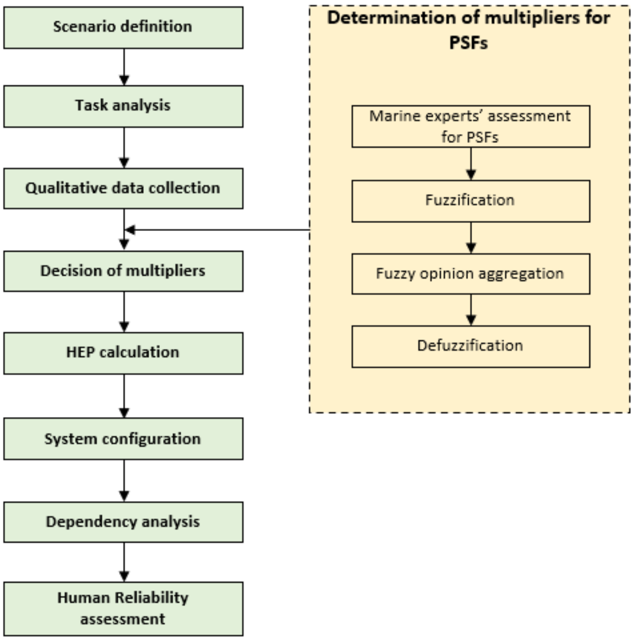

Figure 8. 1 Flow chart of the proposed approach

8.3.1 SPAR-H overview

The Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method was developed to estimate the human error probabilities associated with operator and crew actions and decisions in response to initiating events at commercial U.S. nuclear power plants by Blackman et al. (2008). In the SPAR-H approach, the calculation of HEP rates is straightforward, starting with pre-defined nominal error rates for cognitive versus action-oriented tasks and incorporating performance shaping factor multipliers upon those nominal error rates (Blackman et al., 2008). The SPAR-H method has been applied to human error-related research in various industries. For example, Jahangiri et al. (2016) used the SPAR-H method to analyse and quantify the potential human errors and extract the required measures for reducing the error probabilities in the permit to work system in a chemical plant. In the petroleum industry, the Petro-HRA method, which used the SPAR-H method as the basis for the quantification model, has been developed to analyse human actions as barriers in major accidents and the applicability of human reliability analysis methods (Bye et al., 2017). The Petro-HRA project provides guidance on the comprehensive process of HRA as well as human error quantification. In the maritime industry, Parhizkar et al. (2021) applied the SPAR-H method to estimate the effect of performance shaping factors on the human error probability for the probabilistic risk assessment of decision-making in emergencies of the dynamic positioning drilling unit.

8.3.2 Scenario definition

The scenario defines the scope and boundaries of the analysis and is used as the underlying data for subsequent qualitative and quantitative analysis (Bye et al., 2017). This step focuses on describing the context of individual tasks throughout the whole process. The main objective of scenario development is to create a more detailed description of the event sequence to identify potential human errors better and understand the operational context. The scenario in this paper includes detailed information such as the tasks performed, individuals responsible and their roles, the task location (indicating the working conditions and the external environmental conditions), and the equipment used with their interfaces.

8.3.3 Task analysis

The goal of task analysis in this research can be defined as simply subdividing the functions into tasks, tasks into subtasks, and subtasks into human actions. A task analysis describes the steps performed as part of the activity, providing a method of systematically organising the information collected about the task (Bye et al., 2017). In this paper, two different task analysis methods are utilised. First, a hierarchical task analysis is performed to define the task on the procedure's primary goal, along with subtasks to address the specified duties the operator should complete. Second, the HTA provides a graphical overview of the tasks involved in the analysis scenario. However, hierarchical task analysis is not sufficient to provide appropriate information in the context associated with the tasks. Therefore, a tabular task analysis is utilised to provide more information for experts' judgment and better organise data.

8.3.4 Deriving and rating PSFs

This section begins with the definition of the PSF and describes the step-by-step process of implementing expert evaluations and representing the consensual results in the corresponding PSF multipliers.

8.3.4.1 Define PSFs and guidance for PSF ratings

The selection of PSFs that affect human performance and their assessment criteria should change depending on context. Therefore, the PSFs definition, levels provided by Whaley et al. (2011), were refined by maritime experts with customised guidance to establish and rate characteristics in onboard rescue drills. The provided description of PSFs should be as clear as possible for experts to determine the appropriate PSF rating for the task being analysed while preventing them from selecting the PSF rating mechanically. The criteria for PSFs are set in this section through expert consensus before judgment for PSFs, considering these two opposing aspects simultaneously.

i) Human-machine interface (PSF_1)

The Human-Machine Interface (HMI) PSF refers to the quality of equipment, controls, hardware, software, monitor layout, and the physical workstation layout, where the operator/crew receives information and carries out tasks (Bye et al., 2017). Human-machine interfaces should be appropriately evaluated for the tasks required from two perspectives: interfaces for diagnosis, such as monitors and visible & audible alarms, and interfaces for execution, such as switch buttons, levers, and keyboard(s).

ii) Threat Stress (PSF_2)

Threat stress refers to the expectation or fear of physical or mental harm (Salas et al., 1996). Examples of situations that can cause threat stress on a ship include fear felt by the worker in a confined space, or fear that a lifeboat could fall when it is suspended high above the surface of a davit fall.

iii) Level of experience or training skill (PSF_3)

This PSF refers to the experience and training of the operator(s) involved in the task and should focus on satisfying the experience/skill required by the assigned task, which is identified through task analysis rather than measuring the skills of the worker(s) in a wide range of areas.

iv) Procedures (PSF_4)

The procedures PSF represents the existence and use of formal operational procedures for the task and includes user manuals and instructions for machine and software operations for the task. The procedure is assessed from the following perspectives: whether all required procedures are in place; whether the procedures are easily accessible and visible from the workplace; whether the procedure/manual/instruction contains enough content to perform the task; and whether the content is unambiguous and is easy to understand linguistically and graphically.

v) The complexity of the task (PSF_5)

Task complexity refers to how difficult the task is to perform in the given context. The degree of complexity is measured using different information, including physical and mental hardness, the number of goals, and the number of steps.

vi) The working condition (PSF_6)

The working condition refers to the physical variables in which the work is performed (e.g., temperature, humidity, vibration, noise level, allowable space, and intensity of light). This affects mental state by causing certain moods and emotions. For example, suppose the workplace moves even though it occupies the same space (e.g., a lifeboat). In that case, the different characteristics of lifeboats at stowed positions and lifeboats on the sea should be reflected.

vii) Environmental condition (PSF_7)

Environmental conditions refer to the state of the ship's environment, including weather conditions, sea conditions, and the time of day. The effects of environmental conditions should be evaluated differently depending on the location (e.g., control room, boat on the sea, etc.).

viii) Time pressure (PSF_8)

Time pressure indicates the amount of pressure that requires the operation to be performed in time. Thus, the time pressure depends on the available time compared to the minimum time required. Examples scenarios of negative effects include where: the event to avoid has already occurred or where it is too late to recover within a specified period; a slight delay in time has serious negative consequences; the operator must complete the operation before starting the next sequence operation; and where the task must be performed simultaneously with other tasks or at a specific time.

ix) Ship safety management system (SMS) and supports (PSF_9)

The PSF refers to safety, work, and management support, which consists of three related factors: 1) adequacy of established SMS; 2) the degree of implementation of SMS; and 3) the degree of support offered by the company to perform tasks.

8.3.4.2 Adaption of Fuzzy theory

The selected PSFs have linguistic variables that negatively or positively represent the level of PSFs dealing with expected impacts on performance reliability. In conventional SPAR-H, monolingual variables are determined with 100% faith in the relevant PSF evaluation. However, a limited number of language scales are insufficient to reflect the impact of PSF on human confidence in real-world situations (Ahn and Kurt, 2020). Therefore, fuzzy sets are employed to describe the impact of PSFs better because they offer a useful procedure when dealing with the ambiguity of human error detection problems (Akyuz, 2016). Each PSF connects eight Fuzzy triangular sets to illustrate the impact of each PSF as shown in ***Figure 8. 2***.

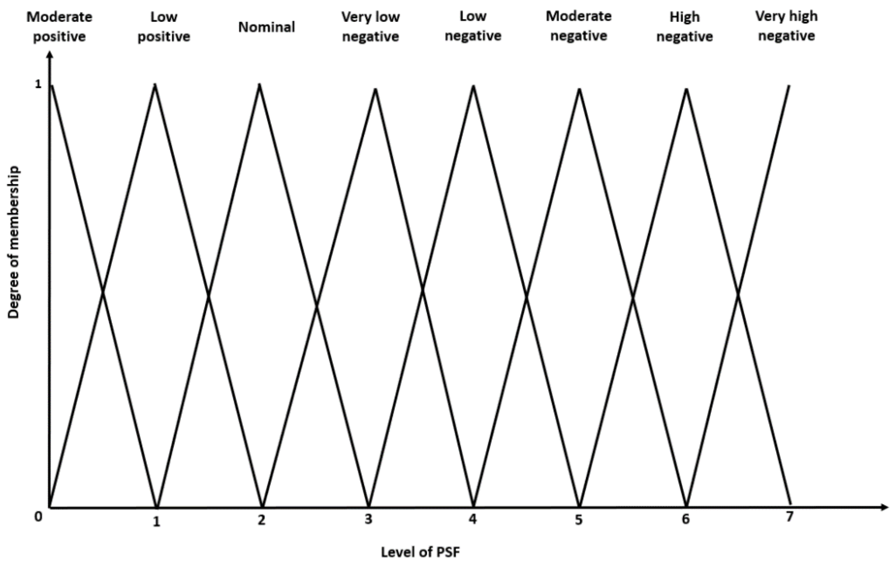

Figure 8. 2 Fuzzy membership for PSFs

The Fuzzy triangular set expressed as (a, b, c) and membership function μ(x) for a linguistic variable is obtained as follows.

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x = b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & \text{Otherwise} \end{cases} \qquad \text{where } a \leq b \leq c \qquad (8\text{-}1)$$

8.3.4.3 Experts' judgment and Fuzzy opinion aggregation

A group of experts are asked to evaluate the level of each PSF considering the characteristics of each task. A linguistic scale for PSF levels and their corresponding Fuzzy set are developed and provided in ***Table 8. 1***.

Table 8. 1 Evaluation of PSFs influence on performance with Fuzzy sets

Level of PSF	Fuzzy set	Number of PSF(N)	SPAR-H multiplier	Interpolated Multiplier
Extremely high negative*	-	-	HEP=1	HEP=1
Very high negative	(6, 7, 7)	N=7	50	50
		6<N<7	-	25N-125
High negative	(5, 6, 7)	N=6	25	25
		5<N<6	-	15N-65
Moderate negative	(4, 5, 6)	N=5	10	10
		4<N<5	-	5N-15
Low negative	(3, 4, 5)	N=4	5	5
		3<N<4	-	3N-7
Very low negative	(2, 3, 4)	N=3	2	2
		2<N<3	-	N-1
Nominal / not applicable	(1, 2, 3)	N=2	1	1
		1<N<2	-	0.5N
Low positive	(0, 1, 2)	N=1	0.5	0.5
		0<N<1	-	0.4N+0.1
Moderate positive	(0, 0, 1)	N=0	0.1	0.1

* If one (or more) PSFs are an extremely high negative case, then the HEP for the corresponding task shall be set to 1 regardless of any other multipliers for the other PSFs.

The purpose of applying the Fuzzy opinion aggregation in ***Figure 8. 1*** is to translate the experts' multiple qualitative assessments of PSF ratings into a single aggregated opinion with Fuzzy opinion and convert it into a crisp value through defuzzification. The modified opinion aggregation procedure, adapted from (Ahn and Kurt, 2020), is made by incorporating a Fuzzy multiple attributive group decision-making methodology by Ölçer and Odabaşi (2005) as follows:

(a) Calculating the degree of agreement (Similarity)

Assume that the Fuzzy set selected by experts A and B as A= (a_1, a_2, a_3), B= (b_1, b_2, b_3), and A and B are standardised Fuzzy sets. Here, S (A, B), which is the degree of similarity between A and B, is measured by equation 6:

$$S(A, B) = 1 - \frac{|a_1 - b_1| + |a_2 - b2_1| + |a_3 - b_3|}{3\ x\ Nmax} \qquad (8\text{-}2)$$

N_{max} is seven because a maximum level of PSF is defined as seven.

(b) Calculating the average degree of agreement (AA)

Let's define AA(Ex_i) as the i_{th} average degree of agreement between $expert_i$ and $expert_j$, and this can be calculated by equation 7:

$$AA(Ex_i) = \frac{1}{D-1} \sum_{\substack{i=1 \\ i \neq j}}^{D} S(Ex_i, Ex_j) \qquad (8\text{-}3)$$

Where D is the number of experts

(c) Calculating the relative degree of agreement (RA)

Let's define RA(Ex_i) as the i-th relative degree of agreement which can be calculated by equation 8:

$$RA(Ex_i) = \frac{AA(Ex_i)}{\sum_{i=1}^{D} AA(Ex_i)} \qquad (8\text{-}4)$$

(d) Calculate the consensus degree coefficient (CC)

Let us define CC(Ex_i) as the consensus degree coefficient for i-th expert, which can be calculated by equation 9:

$$CC(Ex_i) = \beta * w_i + (1 - \beta) * RA(Ex_i) \qquad (8\text{-}5)$$

Where β is a relaxation factor between 0 and 1, note that a Homogeneous group of experts can be calculated by assigning β as 0.

(e) Calculating the aggregation result of the Fuzzy opinion (R_{AG})

The aggregated Fuzzy set R_{AG} can be calculated using the following equation:

$$R_{AG} = \sum_{i=1}^{D} CC(Ex_i) * P(Ex_i) = (S_1, S_2, S_3) \qquad (8\text{-}6)$$

(f) Defuzzification

Finally, an aggregated Fuzzy set R_{AG} for each PSF is converted to a crisp value by a centre of gravity (COG) method, demonstrated below:

Defuzzified rating of PSF = $\frac{S1+S2+S3}{3}$ (8-7)

8.3.4.4 Calculation of multipliers from the defuzzied PSFs ratings

The SPAR-H output in ***Table 8. 1*** provides eight multipliers for eight different PSF ratings, but information on multipliers between integer intervals is not available. To make the multiplier a continuous number, it is assumed that the function follows a linear pattern between adjacent PSF ratings, as shown in ***Figure 8. 3***. The seven functions of linear lines between points can be calculated based on the PSF rating and their multiplier, respectively. The idea of using linear lines was adopted from a research study for dynamic probabilistic risk assessment of decision-making for dynamic positioning drilling units (Parhizkar et al., 2021). For example, a value corresponding to the multiplier for the rating of numbers N between 6 and 7 can be interpolated by using Equation 25N-125. Seven linear functions for each interval are listed in ***Table 8. 1***. This increases sensitivity by providing a corresponding multiplier for consecutive numbers obtained using Fuzzy rather than integers, resulting in an accurate value.

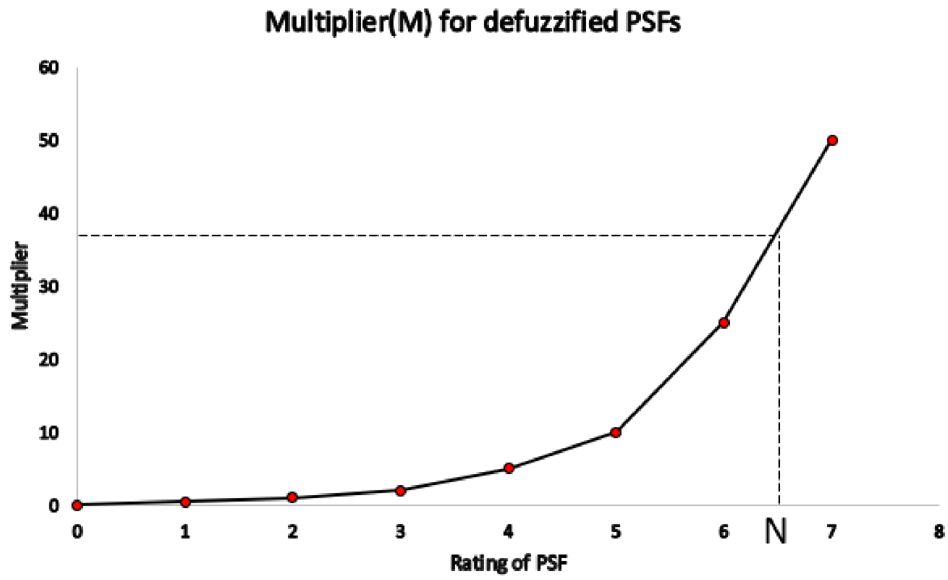

Figure 8. 3 Multiplier for defuzzified PSFs level

8.3.5 Human error quantification

According to SPAR-H, human error probability is calculated as the sum of diagnosis error and execution error. Each error has a nominal failure probability of 0.01 and 0.001, respectively. However, human error probability in this paper is calculated using the below equations, as the 'OR' gate calculation is more reasonable and consistent for computing other failure events than the sum of diagnosis and execution errors.

HEP = 1-(1- Diagnosis Error) x (1- Execution Error) (8-7)

Diagnosis Error = Nominal Diagnosis Error x Composite Multipliers of PSFs (8-8)

Execution Error = Nominal Execution Error x Composite Multipliers of PSFs (8-9)

When there are more than three negative PSFs, human error probability needs to be adjusted by equation (8-10).

$$\text{Adjusted HEP} = \frac{0.01 * \prod multipliers\ of\ PSFs}{0.01 * (\prod multipliers\ of\ \ PSFs - 1) + 1} + \frac{0.001 * \prod multipliers\ of\ PSFs}{0.001 * (\prod multipliers\ of\ \ PSFs - 1) + 1} \quad (8\text{-}10)$$

8.3.6 Modelling of human reliability assessment

Once the human error probability of a sub-task is derived, several rules need to be formulated to calculate the total probability of failure of the entire task. The equations listed in ***Table 8. 2*** are used to obtain the total human reliability in the entire procedure. For sub-task with low or no dependency, failure probability is derived from the multiplication of HEP for sub-task$_i$ in parallel systems. Failure probability is derived from the sum of HEP for sub-task$_i$ in series systems. The minimum HEP of all sub-works is used for parallel subtasks with high or complete dependence. That is, since the task succeeds when any of the sub-tasks is successful, the probability of success of the entire task is assigned to the highest probability of success of the sub-task. The maximum HEP of all sub-works is used for sequential sub-tasks with high or complete dependence. When one of the sub-tasks fails, the highest probability of failure of the sub-task is assigned as the probability of failure of the entire task. The mentioned method proposed by (He et al., 2008) provides a simple and effective way to calculate human reliability, but the following assumptions should be applied. First, all sub-tasks constituting the task should be connected to either parallel or serial systems. Second, the level of dependence on all sub-tasks should be the same within the task. However, since a combination system of series and parallel cannot be applied, there is a limit to its application to actual cases. In addition, the level of dependence on sub-tasks cannot always be assumed to be the same for all cases within the task and may be different. Therefore, this paper proposes a new approach using a Reliability Block Diagram (RBD), assuming each task and sub-task are system components for this HRA modelling. Details will be explained in conjunction with the case study illustrated in Section 8.4.

Table 8. 2 Calculating the Human error probability from HEPs of its sub-tasks (He et al., 2008)

System description	System sub-task dependency	Notation for task HEP & Reliability	
Parallel system	High dependency	$HEP_{Task} = Min\{HEP_{Sub-task\ i}\}$ or	(13)
		$R_{Task} = Max\{R_{Sub-task\ i}\}$	(14)
	Low or no dependency	$HEP_{Task} = \prod(HEP_{Sub-task\ i})$ or	(15)
		$R_{Task} = 1 - \prod(1 - R_{Sub-task\ i})$	(16)
Series system	High dependency	$HEP_{Task} = Max\{HEP_{Sub-task\ i}\}$ or	(17)
		$R_{Task} = Min\{R_{Sub-task\ i}\}$	(18)
	Low or no dependency	$HEP_{Task} = 1 - \prod(1 - HEP_{Sub-task\ i}) \approx \sum(HEP_{Sub-task\ i})$ or	(19)
		$R_{Task} = \prod(R_{Sub-task\ i})$	(20)

8.4 Application of SPAR-H framework to an emergency response drill for man overboard on ships (Case study 3)

Emergency preparedness is of paramount importance in successful emergency responses at sea. Therefore, emergency drills are regularly conducted to maintain acceptable levels of emergency preparedness. However, it needs to be considered that emergency drill operations themselves include significant risks, and there is no evidence that these risks are appropriately considered when planning emergency drill operations. Human error is a main contributor to accidents during emergency drill procedures. The main question posed is how overall risk, including human errors, during an emergency drill can be correctly evaluated. This section demonstrates the application of the new hybrid approach based on the Standardised Plant Analysis Risk Human Reliability Analysis (SPAR-H) method with a Fuzzy multiple attributive group decision-making method to emergency response drills. The method provides a framework for evaluating specific scenarios associated with human errors and identifies contributors that affect human performance. Estimated human errors are utilised to assess human reliability using a new approach based on a system reliability block diagram. For an illustration of the proposed approach, both scenario and procedures for the man overboard rescue drill during ship navigation have been selected because survival crafts were one of the three significant causes of fatality of seafarers, along with entering confined spaces and falling overboard in accordance with the 2001 MAIB report (Ross, 2006). This scenario was developed based on the real shipboard rescue drill observed, and the scenario of a man overboard rescue drill is described in section 8.4.1 to assess PSFs and estimate human error probabilities. The hierarchical task analysis and tabular task analysis are conducted respectively and are described in section 8.4.2.

8.4.1 Scenario definition

The scenario for a rescue boat drill for a man overboard on a ship is described as an illustration of the proposed method. One applied assumption is that the person in charge of watchkeeping is able to observe the man overboard situation. According to SOLAS regulation III/14.1 (IMO, 2018b), the rescue boat should be launched in no more than 5 minutes in such cases.

The ship was under navigation in the open sea from the port of Nagoya to the port of Tokyo for cargo loading. The temperature was 32 °C, and the humidity was 70 %. The wind speed was moderate, and the current speed was relatively high. The vessel was a newly constructed container ship, G/T 6,500 TEU, and the overall vessel condition was described as good. The ship's management company has managed a total of 130 vessels, holding both the company's DOC certificate and SMC certificates for individual ships in effect by an International Safety Management Code (ISM), and has obtained ISO certificates on the quality management system. A month prior, the company conducted an internal audit of the vessel, and two identified nonconformities were rectified. There was a record of the supply of all items on time which were requested by the ship. A total of 22 crew members were on board and were made up of two different nationalities. All crew members have precious experience as rescue crews and have relevant certificates as qualified rescue crew. The captain had more than twenty years of experience and had been working for the company for five years as a captain since the time of ship delivery. The chief officer had seven years of experience and had been onboard the ship for three months. The officer on duty boarded the ship the previous month with three years of experience as a third officer. The watchkeeping crew had five years of experience as AB. Crewman-A, the person in charge of rescue boat control, had seven years of sailing experience as a second officer and had also been on board since ship delivery. Crewman-B and -C had boarded three months ago with three years of experience as oilers. Crewman-D, who was responsible for winch control, boarded six months ago with a twenty-year career as a bosun. As a result of the health record review and interview of all crew members, there were currently no crew members taking medications or experiencing physical and/or mental ill-health. The muster list, including procedures for rescuing men overboard, was posted on the walls of the navigation bridge, each corridor and cafeteria, as well as on their respective duty pocketbooks. The search and rescue procedure, including William’s turn, is visible on the wall of the navigation bridge. The rescue boat davit posts how it works with an illustration. Instruction on the operation of the release hook within the rescue boat is posted with the illustration.

The ship's captain planned to conduct the man overboard drill with a No.1 enclosed type lifeboat, which is assigned as a rescue boat, at 3 p.m. Once the drill begins, the watchkeeping crew on duty identify a man overboard and throws a life buoy. The lifebuoy is a quick-release type, which automatically drops when the safety pin is removed, and the lever is pulled and reports to the officer on duty. The officer on duty reduces the vessel speed, marks the man's location falling into the sea, immediately reports to the captain and the relevant authorities, and uses a public address system and alarm to notify other crew members. As soon as the captain is aware of the situation, they move to the navigation bridge and manoeuvre the ship back to the location where the person was reported as overboard. The rest of the crew gather with personal equipment at the muster station on the boat deck. The chief officer instructs the wearing of personal equipment and performs the inspection. The chief officer then delivers the safety briefing to the crew and instructs them to prepare for the rescue boat launch. These tasks include disconnecting the charging cable socket from the boat, removing the securing wire and lashing stopper, and connecting a painter line to hook on the rescue boat FWD. Crewman-A, -B, and -C are designated as rescue crews and board the rescue boat, while the rest of the crew are responsible for helping to launch and recover a rescue boat by controlling a davit winch. The rescue boat is a davit launching type and is lowered by gravity when the winch brake is released. The winch breaker is operated using a lever after removing the safety pin. The boat is raised by pressing a button on the remote controller to operate the winch. Crewman-D is responsible for the winch operation. The boat release hook system can be operated in both on-load and off-load conditions. A hydrostatic interlock device is installed to prevent crew members from falling out before reaching sea level. The hydrostatic interlock can be manually released if the water pressure is not working properly due to severe sea conditions even though the boat has reached water level. In this case, releasing is called an on-load release, while scenarios where the boat is buoyant and released without force applied to the hook are deemed an off-load release. The release hook is operated using a lever after removing the safety pin by Crewman-A. Then, the pull lever on the rescue boat's FWD is pulled, and the painter line is removed before rescue operations begin. Once the rescue boat is wholly removed from the primary vessel, the rescue boat is manoeuvred to begin rescue operations. At the end of the rescue activity, a suspension link from the davit fall is connected to the hook of the rescue boat for boat recovery. If the rescue boat is properly connected to the wire of the davit, the winch is activated to raise the boat to the stowed position. More specific tasks are described in the task analysis.

8.4.2 Task analysis

The hierarchical task analysis for the procedures of the rescue boat drill for a man overboard is shown in ***Figure 8. 4***. The procedure consists of eight main tasks, which are i) Preparation of rescue activity, ii) Ship manoeuvring, iii) Assembly and safety briefing, iv) Preparation of boat lowering and embarkation, v) Boat testing and lowering, vi) Boat release, vii) Manoeuvring and rescue activity, and viii) Boat recovery. Additional information obtained from the tabular task analysis is described in ***Table 8. 3***.

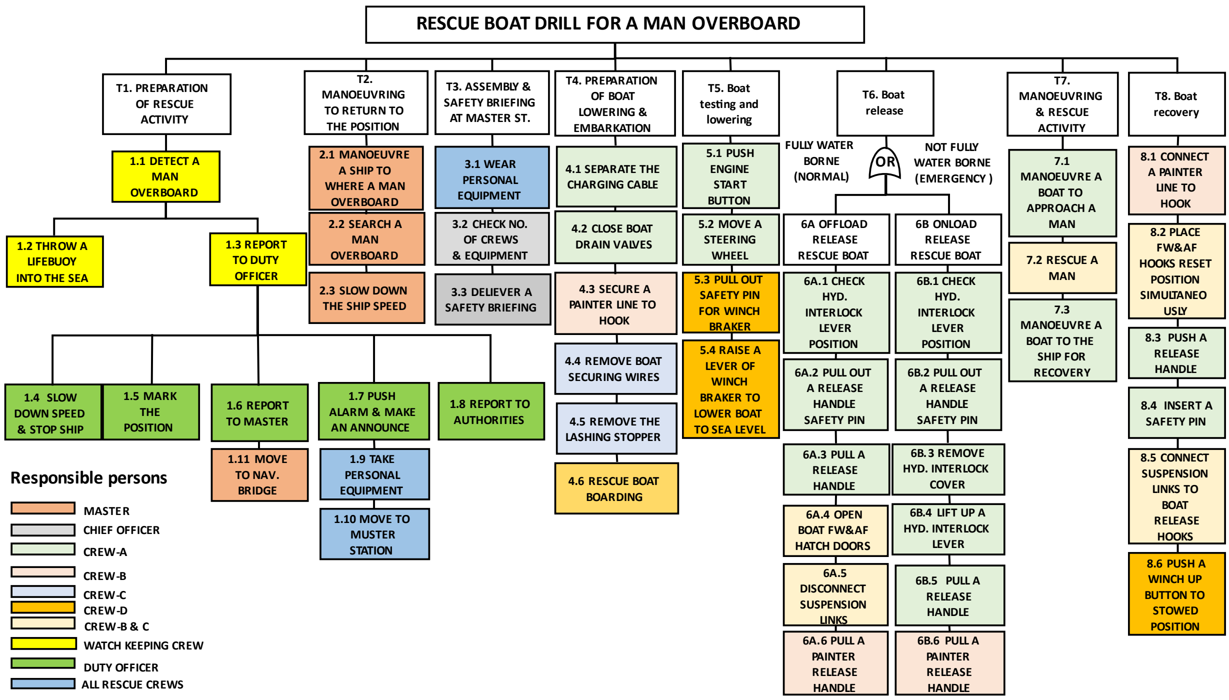

Figure 8. 4 Hierarchical Task analysis

Table 8. 3 Tabular task analysis (part of, Full TTA is listed in Appendix A)

Task	Responsible person	Location	H-M interface	Equipment	Required additional manual
1.1 Detect a man overboard	Watchkeeping crew	Wing bridge	N/A	N/A	N/A
1.2 Throw a lifebuoy into the sea	Watchkeeping crew	Wing bridge	Safety pin, lever handle	Quick-release lifebuoy	Manual for quick release lifebuoy
1.3 Report to the duty officer	Watchkeeping crew	Nav. bridge	N/A	N/A	N/A
1.4 Slow down speed and stop the ship	Duty officer	Nav. bridge	Lever handle	Engine telegraph	N/A
1.5 Mark the position where a man overboard	Duty officer	Nav. bridge	Display screen	Chart or ECDIS	N/A
1.6 Report to Master	Duty officer	Nav. bridge	Telephone	Telephone	N/A
1.7 Push alarm and make an announcement	Duty officer	Nav. bridge	Push-button, Announce device	Alarm system, P.A. system	Manual for alarm system, Manual for PA system
1.8 Report to Authorities	Duty officer	Nav. bridge	VHF	VHF	Manual for radio equipment and contact details
1.9 Take personal equipment	All rescue crews	Cabin room	N/A	Personal equipment	N/A

8.4.3 Performance shaping factors assessment

Five maritime experts are carefully selected for this assessment, where experts are asked to select a linguistic scale of PSFs for each task. Then, qualitative expert opinions are aggregated. Experts' relative importance is considered a heterogeneous group, depending on their background. A relaxation factor (β) is assumed to be 0.4, and relative importance w_i among experts is determined as 0.20, 0.18, 0.21, 0.20 and 0.21 for five experts. As an example, specific opinion aggregation for Task 8.2 is illustrated in ***Table 8. 4*** and ***Table 8. 5***. Once experts' judgment and Fuzzy opinion aggregation are completed, the aggregated rating is converted to the multiplier for each PSF for a human error quantification, following equations in ***Table 8. 1***.

Table 8. 4 Experts' evaluation for PSFs of task 8.2

PSFs	PSF1	PSF2	PSF3	PSF4	PSF5	PSF6	PSF7	PSF8	PSF9
Expert 1	N	N	N	N	N	LN	LN	VLN	N
Expert 2	VLN	N	N	VLN	N	VLN	LN	N	LP
Expert 3	N	N	N	VLN	N	LN	VLN	VLN	N
Expert 4	N	VLN	N	VLN	N	N	N	N	LP
Expert 5	N	VLN	N	VLN	N	LN	LN	N	N
Aggregated rating	2.18	2.39	2.00	2.81	2.00	3.44	3.42	2.40	1.61
Interpolated multiplier	**1.18**	**1.39**	**1.00**	**1.81**	**1.00**	**3.32**	**3.27**	**1.40**	**0.81**

*N is nominal, VLN is very low negative, LN is Low negative, LP is Low positive

Table 8. 5 Opinion aggregation working condition of task 8.2

Degree of agreement(S)		**The relative degree of agreement (RA)**	
S12	0.86	RA(Ex1)	0.21
S23	0.86	RA(Ex2)	0.2
S13	1	RA(Ex3)	0.21
S14	0.71	RA(Ex4)	0.18
S24	0.86	RA(Ex5)	0.21
S34	0.71		
S15	1	**Consensus degree coefficient (CC)**	
S25	0.86	CC(Ex1)	0.2
S35	1	CC(Ex2)	0.19

S45	0.71	CC(Ex3)	0.21
The average degree of agreement (AA)		CC(EX4)	0.19
AA(Ex1)	0.89	CC(Ex5)	0.21
AA(Ex2)	0.86		
AA(Ex3)	0.89	**Result of aggregation (Rag)**	
AA(Ex4)	0.75	(2.44, 3.44, 4.44)	
AA(Ex5)	0.89	**Defuzzified rating**	**3.44**

8.4.4 Human error quantification

The human error probabilities for each sub-task of a man overboard procedure during the whole rescue boat drill are listed in ***Table 8. 6***. Human error probability is computed based on the SPAR-H human quantification technique, which is described in section 8.3.5.

Table 8. 6 Human error probability for each sub-task of a man overboard procedure

Task	Diagnosis Error	Execution Error	Human Error
1. Preparation of rescue activity			
1.1 Detect a man overboard	2.04E-02	2.04E-03	2.24E-02
1.2 Throw a lifebuoy into the sea	5.72E-03	5.72E-04	6.29E-03
1.3 Report to the duty officer	3.60E-03	3.60E-04	3.96E-03
1.4 Slow down speed and stop the ship	3.30E-03	3.30E-04	3.63E-03
1.5 Mark the position where a man overboard	3.05E-03	3.05E-04	3.36E-03
1.6 Report to Master	2.23E-03	2.23E-04	2.45E-03
1.7 Push alarm and make an announcement	2.03E-03	2.03E-04	2.23E-03
1.8 Report to Authorities	9.81E-03	9.81E-04	1.08E-02
1.9 Take personal equipment	2.91E-03	2.91E-04	3.20E-03
1.10 Move to muster station	3.65E-03	3.65E-04	4.01E-03
1.11 Move to Navigation bridge	3.65E-03	3.65E-04	4.01E-03
2. Ship manoeuvring			
2.1 Manoeuvring a ship to where a man overboard	1.59E-02	1.59E-03	1.75E-02
2.2 Search a man overboard	2.08E-02	2.08E-03	2.28E-02
2.3 Slow down the ship speed	1.58E-03	1.58E-04	1.74E-03
3. Assembly and safety briefing			
3.1 Wear personal equipment	3.33E-03	3.33E-04	3.66E-03

3.2 Check number of crews and their equipment	2.84E-03	2.84E-04	3.12E-03
3.3 Deliver a safety briefing	9.52E-03	9.52E-04	1.05E-02
4. Preparation of boat lowering and embarkation			
4.1 Separate the charging cable	6.47E-03	6.47E-04	7.11E-03
4.2 Close boat drain valves	1.12E-02	1.12E-03	1.23E-02
4.3 Secure a painter line to the rescue boat painter hook	1.70E-02	1.70E-03	1.86E-02
4.4 Remove boat securing wires	1.33E-02	1.33E-03	1.47E-02
4.5 Remove the lashing stopper on boat davit	1.33E-02	1.33E-03	1.46E-02
4.6 Rescue boat boarding	8.24E-03	8.24E-04	9.06E-03
5. Boat testing and lowering			
5.1 Push the engine start button	1.04E-02	1.04E-03	1.15E-02
5.2 Move the steering wheel	8.93E-03	8.93E-04	9.82E-03
5.3 Pull out the safety pin for the winch brake	5.45E-03	5.45E-04	6.00E-03
5.4 Raise the lever of the winch brake to lower the boat to the sea level	1.30E-02	1.30E-03	1.43E-02
6. Boat release			
6A Off-load release rescue boat			
6A.1Check hydrostatic interlock lever position	1.56E+01	1.56E-01	1.56E-02
6A.2 pull out the release handle safety pin	1.04E+01	1.04E-01	1.04E-02
6A.3 Pull the release hand	1.04E+01	1.04E-01	1.04E-02
6A.4 Open the boat F & A hatch doors	5.90E+00	5.90E-02	5.90E-03
6A.5 Disconnect suspension links from hook	2.19E+01	2.19E-01	2.19E-02
6A.6 Pull the painter release handle	1.57E+01	1.57E-01	1.57E-02
6B On- load release rescue boat			
6B.1Check hydrostatic interlock lever position	1.56E+01	1.56E-01	1.56E-02
6B.2 Pull out the release handle safety pin	1.04E+01	1.04E-01	1.04E-02
6B.3 Remove the hydrostatic interlock cover	1.29E+01	1.29E-01	1.29E-02
6B.4 Lift up the hydrostatic interlock lever	1.03E+01	1.03E-01	1.03E-02
6B.5 Pull the release handle	1.04E+01	1.04E-01	1.04E-02
6B.6 Pull the painter release handle	1.57E+01	1.57E-01	1.57E-02

7. Manoeuvring and rescue activity			
7.1 Manoeuvring the rescue boat to approach a man	2.49E-01	2.49E-02	2.68E-01
7.2 Rescue a man (Pull up a man to a boat)	2.65E-01	2.65E-02	2.85E-01
7.3 Manoeuvring the rescue boat to the ship for recovery	2.09E-01	2.09E-02	2.26E-01
8. Boat recovery			
8.1 Connect a painter line to the painter hook	1.67E-01	1.67E-02	1.81E-01
8.2 Place FWD & AFT hooks reset position simultaneously	3.65E-01	3.65E-02	3.89E-01
8.3 Push the release handle	8.98E-02	8.98E-03	9.80E-02
8.4 Insert the release handle safety pin	9.02E-02	9.02E-03	9.84E-02
8.5 Connect suspension links to boat release hooks	2.10E-01	2.10E-02	2.26E-01
8.6 Push the winch up button to stowed position	7.03E-03	7.03E-04	7.73E-03

8.4.5 Human reliability assessment

Once human error probabilities for each sub-task are derived, the final step is to incorporate the human error probability of the sub-task into Hierarchical task analysis in ***Figure 8. 4*** to derive single failure probability for overall assessment by considering system description and dependency of sub-tasks based on the Rules in ***Table 8. 2***. However, as mentioned in section 3.5, to apply these rules to each major task, the sub-tasks should be sequentially connected for evaluation of dependency, and the tasks should also be decomposed to the level to which the same system can be assumed. This section introduces the following techniques for converting from HTA to reliability block diagrams for each task with different characteristics. According to the HTA in ***Figure 8. 4***, the entire procedure consists of eight major tasks, and each task consists of each sub-task. Task 1 consists of eleven sub-tasks that must be successfully performed to complete the preparation of rescue activities. However, sub-tasks for task 1 do not always occur sequentially, and some sub-tasks are linked to multiple sub-tasks. This makes it challenging to connect sub-tasks to either serial or parallel systems for task 1. For example, sub-task 1.1 in ***Figure 8. 4*** requires actions in sub-tasks 1.2 and 1.3, and sub-task 1.3 initiates five actions in sub-tasks 1.4, 1.5, 1.6, 1.7, and 1.8, which may occur in any order. For human reliability modelling, task 1 is decomposed into six different groups of sub-tasks for missions. Each group is assessed by their system configuration and dependency of sub-tasks, as shown in ***Figure 8. 5***. Specifically, the reliability of group R1-1 is assigned as the maximum value of two sub-tasks because sub-tasks 1.1 and 1.2 are high dependencies in the serial system. If one of the two sub-works fails, the mission for group R1-1 fails, and the success or failure of preceding sub-task 1 affects the conditional probability of sub-task 1.2. Similarly, missions from groups R1-2 to R1-6 are configured, as shown in ***Figure 8. 6***. Once reliabilities for all missions from group R1-1 to R1-6 are derived, total reliability for task 1 is assigned as a multiplication of each group's reliabilities, and the value is 8.83E-01 because if any of the six missions fails the task 1 will fail (series system) and six missions have a low dependency. For task 2, reliability for task 2 is assigned as minimum reliability since sub-tasks from 2.1 to 2.3 are in the series system, and three subtasks are a high dependency. For task 3, sub-task 3.2 'Check the number of crews and their equipment is redundancy for the sub-task3.1 'Wear personal equipment relation. It means if any of the two sub-tasks succeed, then task 3 will succeed. Therefore, group reliability for sub-task 3.1 & 3.2 is calculated by equation 16 because they are in the parallel system with low dependency. The total reliability for task 3 is assigned as multiplication of reliabilities for a group of sub-tasks 3.1 & 3.2 and sub-task 3.3 since they are in a series system with low dependency and the value is 9.89E-01 as shown in ***Figure 8. 7***. In task 4, the sub-task from 4.1 to 4.6 is configured in a series system with no dependence. This means that task 4 succeeds only when all sub-works are successfully completed, but the success or failure of each sub-task does not affect the success or failure of other sub-tasks. For example, sub-task 4.1 'Separate the charging cable 'and sub-task 4.2 'Close boat drain vales 'should be completed successfully before rescue boat launching, but each sub-tasks 4.1 and 4.2 does not affect each other's result. The reliability of task 4 is the product of the reliabilities of individual sub-tasks. For task 5, four sub-tasks are classified as two different series systems with high dependency. Then two groups are combined in a serial system with no dependency, as shown in ***Figure 8. 7***. The value of the reliability of task 5 is 9.74E-01. Tasks 6A, 6B, and 7 are all serial sequenced systems with high dependency. Therefore, the minimum reliability of their sub-tasks is assigned as each task's reliability. The values are 9.78E-01, 9.84E-01 and 7.15E-01, respectively. Similarly, the reliability of task 8 is assigned as 6.016E-01 in ***Figure 8. 7***. Finally, the total reliability of the rescue drill scenario can be derived from the reliability block diagram in ***Figure 8. 8*** by computing the reliabilities of each task. In order to determine the final failure probability value for the man overboard drill, eight main tasks should succeed individually. These tasks should be conducted sequentially and highly dependent. Therefore, the minimum reliability of eight tasks is assigned as overall reliability for the whole procedure, and the value is 6.06E-01.

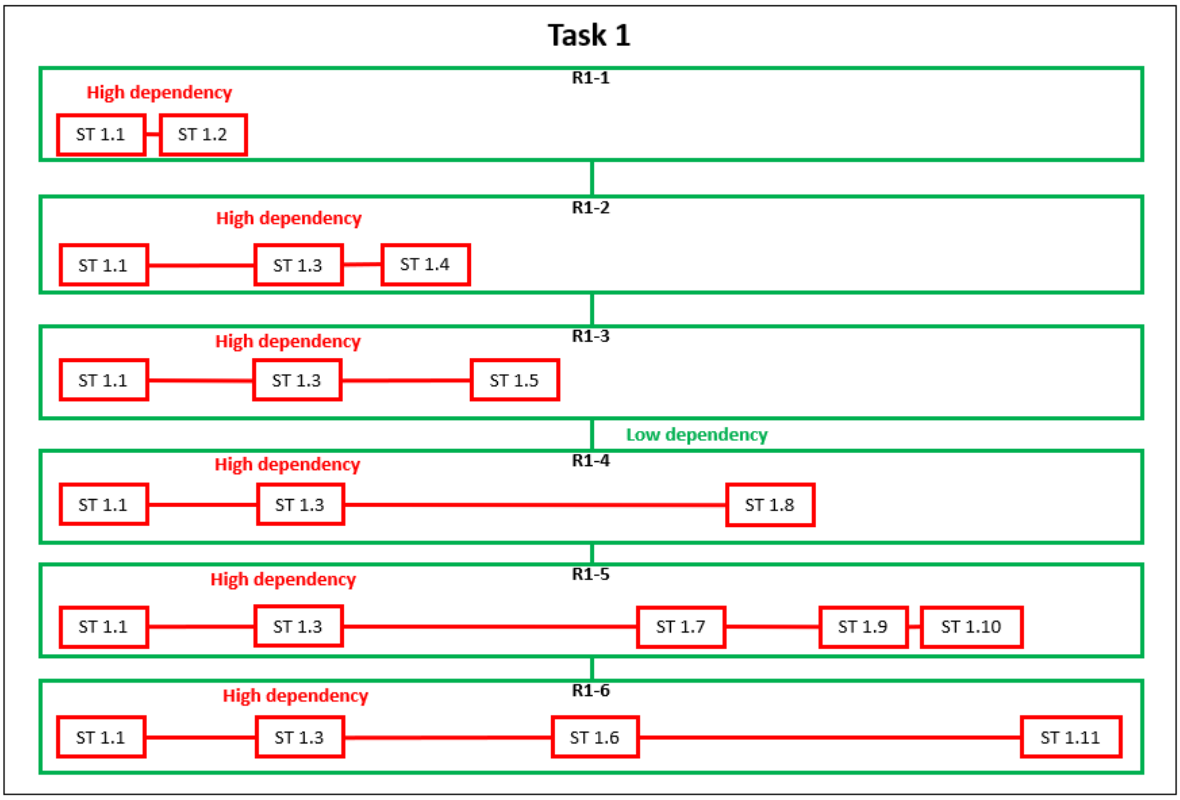

Figure 8. 5 Sub-task level reliability block diagram for task 1

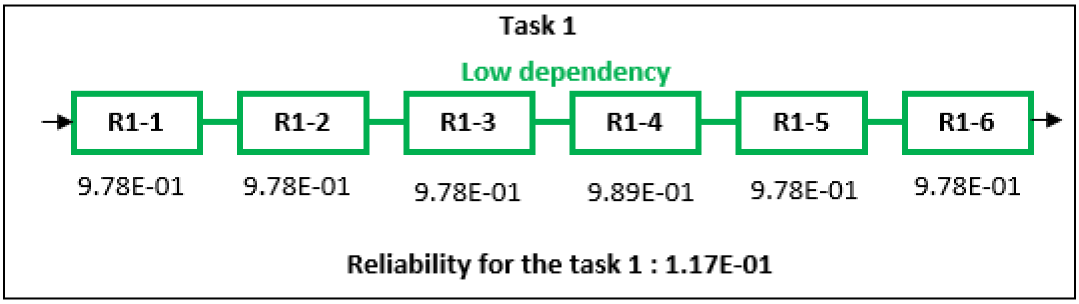

Figure 8. 6 Reliability block diagram for task 1

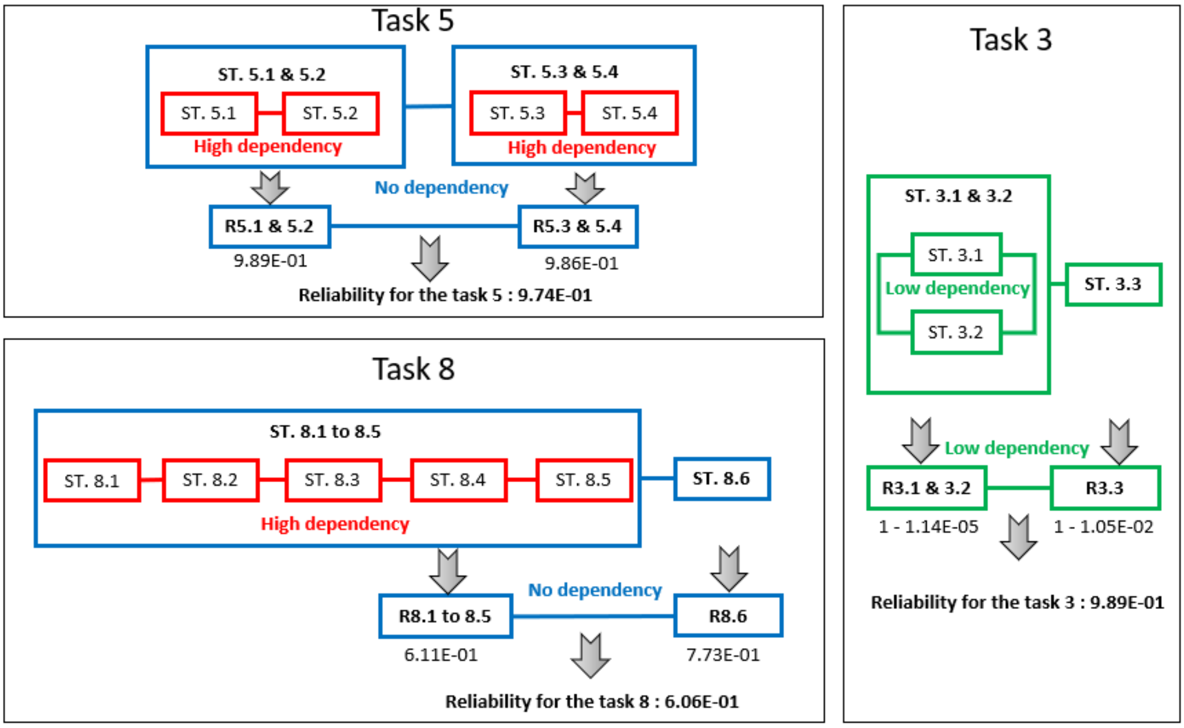

Figure 8. 7 Reliability block diagram for different types of tasks

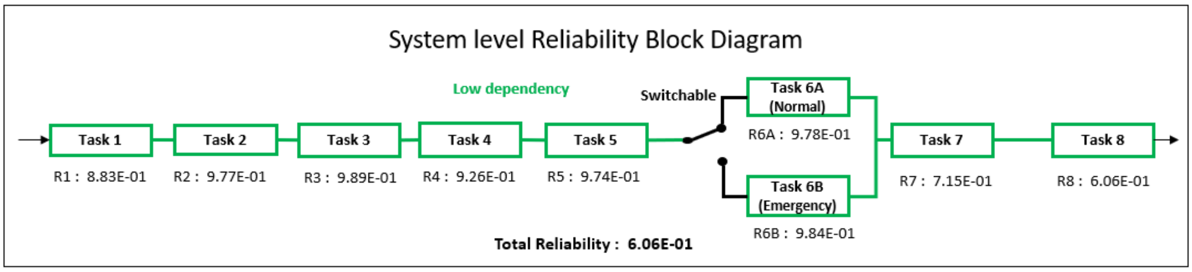

Figure 8. 8 System-level reliability block diagram

8.5 Findings and discussion

The proposed approach presents individual human error probabilities obtained by a proposed method based on a particular maritime scenario: a rescue boat drill for a man overboard procedure. The overall high human error has been demonstrated during the manoeuvring and rescue activity of task 7. This seems to have adversely affected human performance. The narrow space in the enclosed boat restricted the crew's free movement, affected the crew's vision, and even the high noise interrupted communication. Additionally, waves and currents at sea affected the main vessel, making it difficult to control the ship. Therefore, these ambient factors increase stress exerted on the crew and pressure on time for direct rescue

activities. Sub-task 8.2 (Place FWD & AFT hooks reset position simultaneously) in the boat recovery (task 8) has the highest HEP value as of 3.89E-01 during the whole procedure due to the unfavourable circumstances (such as the manoeuvring and rescue activity phase) and each crew member should perform their task simultaneously, even if it is a simple task. The second-highest human error probability is that of sub-task 7.2 (Rescue a man) at 2.85E-01. Sub-task 7.1 (Manoeuvring the rescue boat to approach a man) at 2.68E-01; sub-task 7.3 (Manoeuvring the rescue boat to the ship for recovery) & sub-task 8.5 (Connect suspension links to boat release hooks) are 2.26E-01 in order, are also at notably high probabilities of failure. Conversely, sub-tasks 2.3 (Slow down the ship speed), 1.6 (Report to Master) and 1.7 (Push alarm and make an announcement) show the lowest HEP with a range from 1.74E-03 to 2.45E-03. In human reliability assessment for each task, tasks 1 through 4 are to prepare for rescue operations where human reliabilities intervals range from 8.83E-01 to 9.89E-01. The operations in the rescue boat are divisible: task 5 for embarkation at the boat stowed position; task 6 for boat release where the boat is hung on the wire fall; task 7 for rescue activities; and task 8 for boat recovery at sea condition. The human reliability ranges for tasks 5 through 8 in a rescue boat operation are in 6.06E-01 to 9.84E-01. These different human reliability intervals, occurring in the same rescue boat and while performing similar tasks, demonstrate that even if the task takes place in the same workspace with the same crew, human performance changes depending on the characteristics of the circumstance. The task with the lowest reliability is the process of boat recovery of task 8. Compared to task 6(Release rescue boat) 's reliability interval of 9.78E-01 to 9.8.4E-01, depending on the type of boat release, the boat recovery process causes more human errors than release.
Interestingly, whether the rescue boat is released on-load or off-load condition, it does not change the reliability of the entire process. This means tasks that take place in sea conditions after a rescue boat is released from a wire fall significantly affect whole reliability. Finally, the PSFs contributing to human error in rescue scenarios for each task phase are illustrated in ***Figure 8. 9***. However, the extent to which PSF affects human performance may vary from sub-task to sub-task, even within the same task. The most significant contributing factors are the ambient conditions for the workplace and environmental conditions in tasks 6 to 8 related to the rescue boat operation at sea, which is at the core of the procedures. These findings reveal the impact of any task on the operation's success or failure, allowing for the adoption of

additional safety measures for the indicated task. Furthermore, assessments of elements that affect human performance contribute to the efficient improvement of system safety.

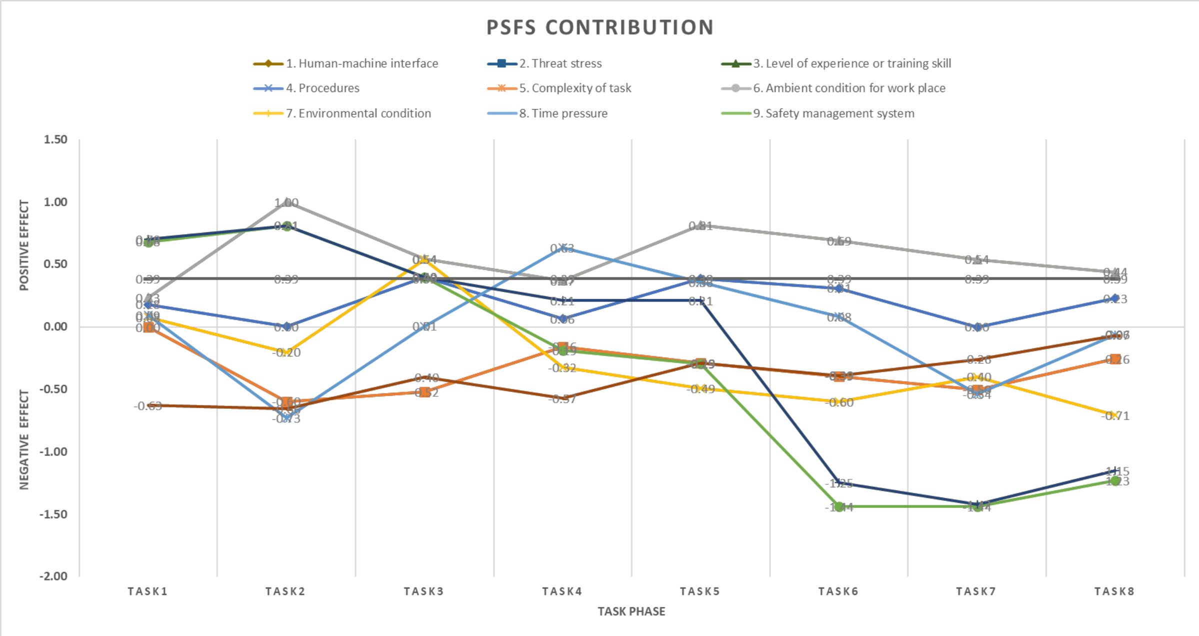

Figure 8. 9 PSFs contribution for task phases

8.6 Chapter summary

Human error is one of the main contributing factors to failure during emergency preparedness in maritime transportation. Although human error assessment is a critical issue in maritime safety, the quantification process is difficult due to the limited human error data. Therefore, many parts of the quantification process rely on the qualitative judgment of experts. Therefore, forming consensus properly and effectively quantifying the collected diverse opinions is important. In this context, this paper introduces a new framework-based SPAR-H approach applicable to maritime emergency drill scenarios and illustrates practical rescue boat drill procedures for a man overboard. There are various characteristics and expected advantages of the proposed method. Firstly, the proposed method provides refined PSFs with customised guidance to reflect onboard rescue drills' characteristics for expert opinion collection. The selected PSFs have a language variable representing the level of PSF that deals negatively or positively with the expected impact on performance reliability and developed a Fuzzy set to better describe the impact of PSF. Secondly, the Fuzzy opinion aggregation method converts experts' multiple qualitative assessments of PSF ratings into one integrated opinion with Fuzzy opinions and converts them into crisp values through defuzzification. SPAR-H then calculates human error probabilities based on PSF ratings from an expert's opinion aggregation method for tasks during rescue boat drills. This hybrid approach may enhance the reliability and consistency of the outcomes. Notably, the novel approach to model a human reliability assessment from individual human error probabilities based on a reliability block diagram applies to various systems. In addition, this approach effectively displays the relationship between complex tasks in a simplified way, which cannot be achieved solely by applying hierarchical task analysis.

In conclusion, maintaining emergency preparedness is undeniably essential in ship operation. However, it is also important to evaluate whether training for emergency preparedness is in a suitable state to be implemented. This study provides a framework for conducting a human reliability assessment of emergency training, which can help ship operators in the decision-making process and positively impact the safety of ship operations and maritime safety.

9 System reliability assessment in the modern complex system

9.1 Chapter overview

By introducing autonomous or software-controlled systems, human operators are increasingly required to perform cognitive-intensive tasks in addition to existing labour-intensive tasks. Thus, it will be more difficult to identify human roles in future complex systems with traditional approaches such as hierarchical task analysis used in the conventional HRA. This chapter demonstrates a novel systematic approach for a human reliability assessment to better understand human activities in complex systems. The proposed framework is a hybrid method combining the System Theoretic Process Analysis (STPA) and the Success Likelihood Index Method (SLIM) to assess the system reliability. STPA is adopted to analyse the interaction relationship between different types of system components. The primary purpose of STPA is to find and analyse human activities that affect the risk contained in human-machine interaction systems. Then the identified human activities are evaluated and quantified by the SLIM as a probability of human error. Finally, the system reliability block diagram represents the derived human error probabilities to assess the entire system for a probabilistic risk assessment. To accomplish this goal, the chapter is organised in the following manner: Section 9.2 introduces research motivation and LNG bunkering overview. Section 9.3 describes the proposed method, and section 9.4 presents a case study of the Emergency Shutdown system during LNG ship-to-ship bunkering. The findings and discussion are presented in Section 9.5, followed by a chapter summary in Section 9.6.

9.2 Research motivation and background

The development of new technologies in the maritime industry brings about a drastic change in how the maritime industry approaches new challenges and opportunities. Many ship operations have been changing with state-of-the-art technology to partially or fully automated control systems. Due to these changes, human roles in cognitive-intensive behaviour are becoming increasingly crucial in maritime operations in addition to the existing labour-intensive behaviour. In terms of safety management, it will be more difficult to identify human roles with traditional approaches such as hierarchical task analysis commonly used in the traditional Human Reliability Assessment (HRA) in future complex systems. Traditionally, hazard analysis techniques such as failure tree analysis (FTA) and failure mode and impact analysis (FMEA) have been in widespread use for decades. However, conventional approaches are not suitable for capturing the effects of changes in modern, more complex systems that are software-intensive and have a socio-technological component. In particular, techniques for identifying and quantifying the rapidly changing human roles are not adequately considered in human reliability analysis. Therefore, it is necessary to examine what improvements have been made to apply the existing HRA techniques to the changed or added human role. In this context, this chapter proposes utilising a novel systematic approach for a human reliability assessment in a human-machine interaction system. In the field of hazard analysis, the System Theoretic Process Analysis (STPA) is considered to be a relatively new technique that is based on the System-Theoretic Accident Model and Processes (STAMP) (Leveson and Thomas, 2018). On the other hand, the Success Likelihood Index Method (SLIM) is an HRA technique for determining the likelihood of human error while completing a specific task (Embrey et al., 1984). In this study, the System Theoretic Process Analysis (STPA) is employed to identify human roles and defective interactions between different types of system components in a complex system. The SLIM is embedded in human error quantification to be incorporated into probabilistic risk assessment. For an illustration of this new approach for a complex system in maritime operation, the emergency shutdown system for the LNG ship-to-ship bunkering process is adopted because the emergency response through the human-machine interaction during safety-critical operations like LNG bunkering should be carefully evaluated in terms of safety to prevent loss of life, environmental pollution, and damage to property.

9.3 Methodology

The suggested framework is a hybrid method for assessing system reliability that combines the STPA and the SLIM. As new technologies, like autonomous ships and software control systems, have been integrated into maritime operations, human responsibilities in the maritime system have shifted away from labour-intensive behaviours to cognitive demand intensive activities. As a result, it is vital to assess what enhancements have been made to the current HRAs to adapt them to new human roles. Therefore, systematic approaches are proposed to better understand human activities in the complex system.

In this context, the STPA is adopted to analyse the interactions between various system components (i.e., humans, software, and machines). The fundamental objective of applying STPA in this study is to identify human actions that influence the system risk. The SLIM evaluates and quantifies the identified human activities as a form of human error probability. The computed human error probabilities are then utilised to assess system reliability through a system reliability block diagram for a probabilistic risk assessment. Meanwhile, unsafe control actions discovered by STPA can be studied further, as needed by the study scope, to identify potential loss scenarios, develop further requirements, identify mitigations, and make safety recommendations. This technique enables qualitative or quantitative assessment of system reliability, or a combination of the two, depending on the analysis objective. The suggested method's flow chart is depicted in ***Figure 9. 1***.

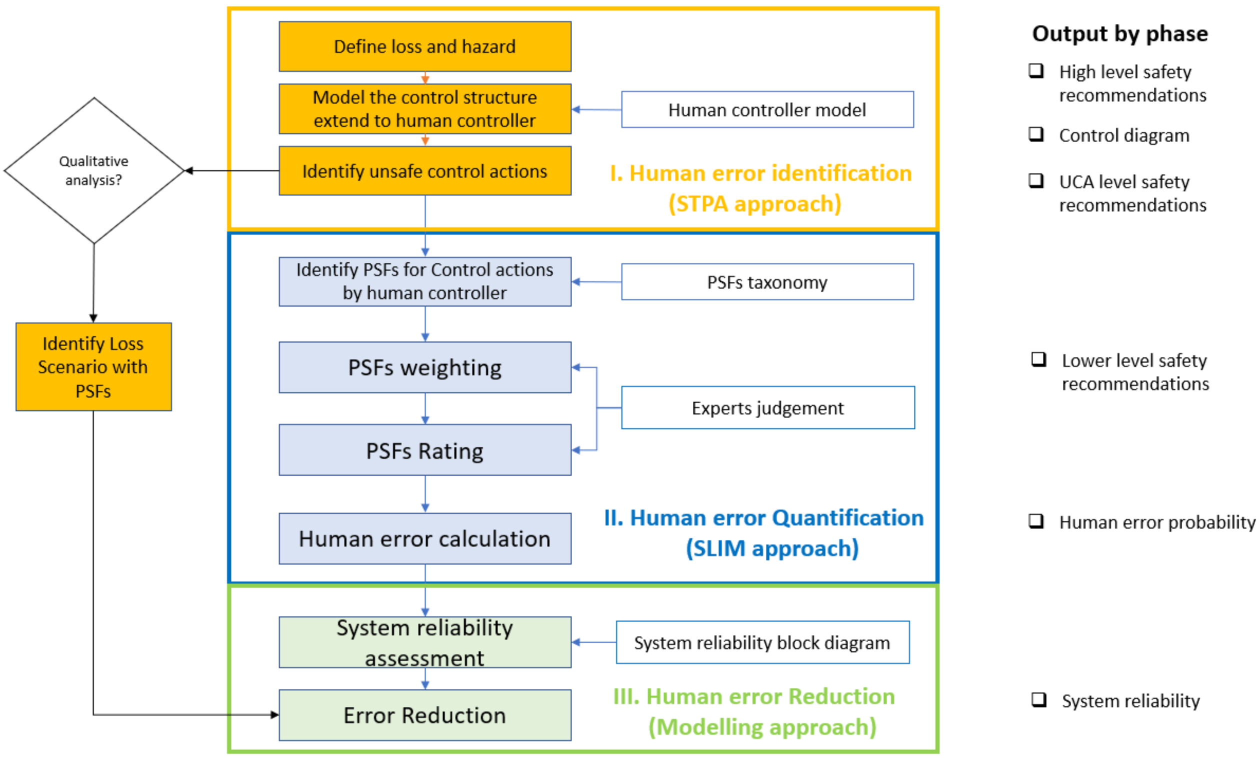

Figure 9. 1 Flow chart of the proposed approach

9.3.1 Systematic approach

The STPA is divided into four distinct stages (Leveson and Thomas, 2018). The first stage is to establish the analysis's aim. This stage addresses the loss, system's description, and boundaries. The second stage is to build what is known as a control structure, which is a model of the system. The control structure describes the system as a feedback and control loop representing functional linkages and interactions. Typically, the control structure begins at an abstract level and is adjusted repeatedly to incorporate more detailed knowledge about the system. The third step is to analyse the control action to determine how it can result in the loss. The discovered unsafe control actions are then utilised to define the system's functional needs and restrictions. Finally, the last phase determines the reasons for unsafe control in the system. The following section 9.4 will discuss the details of these four steps with a case study.

9.3.1.1 Modelling the human controller

Humans can be viewed as a system component from a system perspective. In comparison to logical computer controllers, human decision-making processes, on the other hand, are difficult to predict and far more complicated. Additionally, an external component, Performance Shaping Factors (PSF), which the system cannot regulate, influences the human controller's reaction. Thus, when anticipating the human controller's response in a particular environment, the process model for the human controller should be more thorough and enlarged to incorporate external factors affecting human performance. In this context, a novel process model for human controllers was presented, as seen in ***Figure 9. 2***, to explain human errors and identify contributing factors more efficiently. The model of the human controller is divided into two components: diagnosis and execution. It contains PSFs that affect human performance. The diagnosis is further divided into the 'sensation & perception' process to know updated system conditions and the 'decision-making' process to decide how to act. Sensation and perception are separate processes but are very closely related. Therefore, even though it is challenging to distinguish processes accurately, it is helpful to identify specific PSFs affecting each process. For example, if a fire alarm occurs, the volume of the sound may affect the sensation. However, additional identification methods such as the type and interval of the sound are required to recognise that it is an alarm for fire.

In this study, system inputs representing the beliefs about process are classified into process variables, process environment, and process behaviour. The process variable is the value of a specific part of a current measured process, such as speed or pressure. The process environment is defined as an event that is not directly controlled by the system controller but affects the process. The last factor, process behaviour, indicates responsibility for the control action. Process behaviour can be expressed in control modes such as automatic and manual modes. For these three factors, a loss scenario can be created by considering the factors affecting each process of sensation and perception. Diagnosis's following process consists of making decisions about action objects, action types, and action sequences based on the updated system information. When the diagnosis process is over, the next execution stage appears as physical behaviour. Each box in the human model is useful for identifying specific performance shaping factors that affect the human controller's thoughts and behaviours.

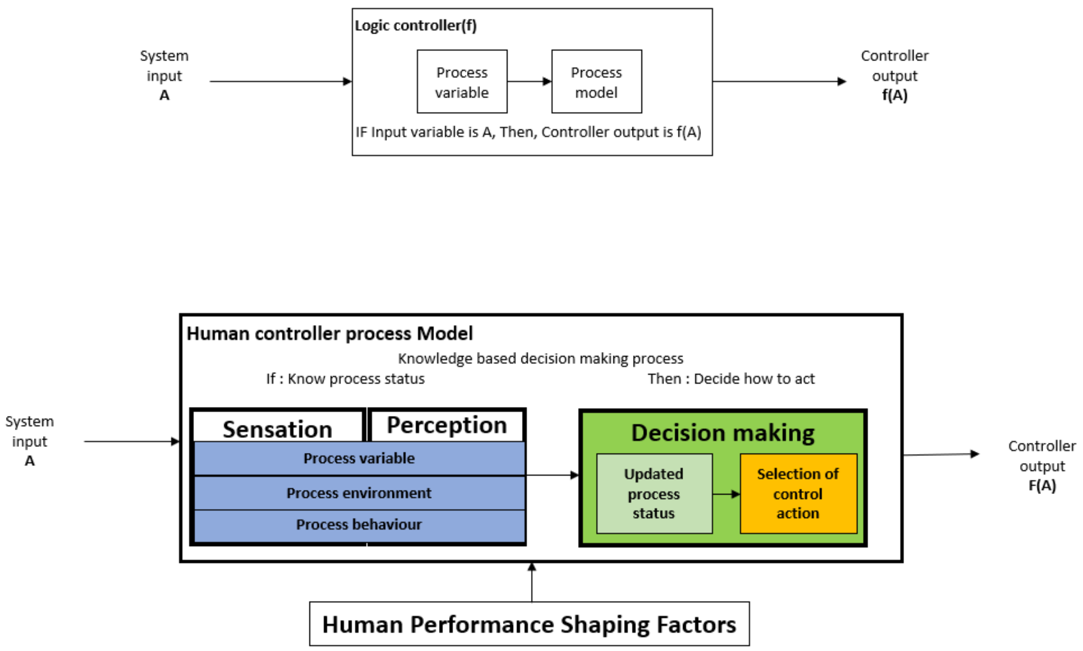

Figure 9. 2 Logic controller VS Human controller model

9.3.1.2 PSFs derivation

Although numerous studies have been undertaken on PSF, there is still considerable uncertainty regarding its effect on human performance. Additionally, expert opinion is employed to establish the relative relevance and ranking of PSF. However, the evaluation results vary significantly depending on each expert group. How PSF affects human performance is frequently mischaracterised, and evaluations are commonly conducted without using suitable criteria. ***Table 9. 1*** shows that the evaluation criteria for customised PSFs were created to mitigate this issue. The presented PSFs are derived from a review of the literature for existing HRA methods such as HEART (Williams, 1985), THERP (Swain, 1964), CREAM (Hollnagel, 1998) and SPAR-H(Blackman et al., 2008). The most effective strategy to cope with uncertainty is to break it down to a level we can easily understand. In addition, when the PSF of the application form is determined, the rating of PSF is measured as an evaluation criterion. For example, if the input interface is applied with a sound alarm, the rating is evaluated according to the degree of audibility. In the SLIM approach, the PSF is utilised to compute the human error. The criteria are also used as keywords to identify the loss scenario of STPA. Therefore, the developed PSF taxonomy can be used for quantitative and qualitative analysis.

Table 9. 1 Customised PSFs for ESD system during LNG ship-to-ship bunkering operation

PSFs	**Possible type of application**	**Keywords**
1. Interface (Input device)	Monitor screen, Digital number display, Analogue gauge, Visible alarm (Lamp), Audible alarm, Other	Layout, Visibility, Distinctness (unclear marking), Consistency, Audibility, Aesthetics, Malfunction, Other
2. Interface (Output device)	Keyboard, Mouse, Handle, Lever, Pushbutton, Switch, Touch surface, Other	Usability, Layout, Distinctness (unclear marking), Consistency, Malfunction, Accessibility, Other
3. Procedure	Operations manual, Emergency manual, Maintenance manual, step by step procedure, Checklist, Graphic display, Diagram, Other	Missing, Unclear, Vagueness, Wrong, Complex, Other
4. Working condition	Stress, Ambient condition, Facility, Lighting, Noise, Interruption,	Bad, Hot, Noisy, Dark, Bright, High, Low, Other

	Temperature, Humidity, Fatigue, Other	
5. Time	Available time, Required time, Other	Shortage, enough, Other
6. Experience & training	Relevant Experience, Formal Training, Technical knowledge, Skill, Qualification, Emergency response training, Other	Lack, Short, Not done, Unfamiliar, Not enough, Other
7. Environmental condition	Weather, Sea condition, Wind, Time of day, Ship movement, Visibility, Interference from other vessel's navigation, Mooring condition, other	Bad, Windy, Tough, Rain, Snow, Slippery, Dark, Late, Other
8. Complexity	SIMOPS, Workload, Working time, Number of goals, Designated person, Other	Over, Many, Long, Not assigned, Hard, Other
9. Organisational factors	Communication, Safety culture, Supervision, Safety Management System, Human resource, Supply, Audit, Monitoring, Other	Lack, not enough, Not done, Difficult, Low, Not efficient, Wrong, Other

9.3.2 Human error quantification with SLIM

The Success Likelihood Index Method (SLIM) is an evaluation tool for human reliability used to quantify the likelihood of human error when completing a specific duty (Embrey et al., 1984). It is a practical and straightforward method to estimate human error when obtaining human error data is difficult (Park and in Lee, 2008). Performance Shaping Factors (PSF), which have a significant impact on human performance, are quantified in SLIM and changed to a preference index form (Akyuz, 2016), allowing for the quantitative representation of external factors impacting human performance in the form of human error. SLIM consists of a six-step process that includes the following steps: 1) task analysis and scenario definition, 2) PSF derivation, 3) PSF weighting, 4) PSF rating, 5) SLI calculation, and 6) SLI to HEP conversion. In this paper, since the STPA, as mentioned above, is adopted to identify human tasks in the complex system in more detail, the SLIM method is used for the rest of the steps except for the first step.

9.3.2.1 PSF weighting

Prior to assigning a rating to each PSF, its relative importance should be determined, as not all PSFs have the same effect on human performance. Additionally, to accurately reflect the features and characteristics of each task, the relative importance of PSFs should be measured for each task independently. Experts evaluate the significance of each PSF on a scale of 0 to 100 and then determine the mean weight value. As in equation (9-1), the normalised weight is generated by dividing the mean weight value by the sum of the mean weights.

$$\text{Normalised Weight}(W_i) = \frac{Mean\ Weighti}{\sum_{i=1}^{9} Mean\ Weighti} \quad (9\text{-}1)$$

9.3.2.2 PSF rating

PSF rating refers to the expert judgment process determining how each PSF impacts each task. Selected experts assign a score to each PSF ranging from 0 to 100. To minimise the deviation of expert evaluations, the Likert scale and upper and lower bounds for the relevant ratings are provided in ***Table 9. 2***. According to evaluation criteria such as professional, service time, and experience, the selected expert group has relative importance (w_j). The consensus rating (R_i) for each PSF_i is computed as the sum of the values obtained by multiplying the j-th expert's rating for the i-th PSF (R_{ij}) by the expert's relative importance (w_j) as specified in Equation (9-2). Note that since w_j is a normalised value for all experts, a separate normalisation process is not required for the consensus rating.

Consensus rating (R_i) for PSF_i = $\sum_{j=1}^{n} w_j \times R_{ij}$ (9-2)

Table 9. 2 Likert scale and score bounds

Likert scale	Upper bound	Lower bound
Highly positive	85	100
Positive	70	85
Moderate	50	70
Negative	30	50
Highly negative	0	30

9.3.2.3 Human error calculation

After the consensus rating R_i and normalised weight W_i of PSF_i are determined, the Success likelihood Index (SLI) for each task is computed by equation (9-3).

$SLI = \sum_{i=1}^{9} \mathrm{Ri\ x\ Wi}$ (9-3)

Accordingly, the SLI value is converted into the HEP value using equations (9-4), where a and b are constant (Embrey et al., 1984). The details of constant determination will be discussed in the following section 9.4.

Log of Probability of Success = a**SLI* + b (9-4)

9.3.3 Modelling system reliability

After deriving the probability of nominal failure, including human errors for each component, the method of integrating the system reliability of each task into the probability risk model should be considered. Therefore, this section adopted an approach by Ahn et al. (2022) using a Reliability Block Diagram (RBD), assuming each human task is a system component for reliability modelling. In addition to human tasks, events, functional elements, and any behaviour expressed as success and failure can be regarded as elements of the system. The system configuration method and dependency between components must be defined to model a reliability block diagram. For system configuration, if a sub-system is essential for the mission success of the overall system, it should be modelled as a series component. The parallel configuration indicates that the primary function of that sub-system is duplicated, thus allowing a switch over to the redundancy in the event of failure. The example of each system is illustrated in ***Figure 9. 3***. The next step is to determine the dependence between components. Dependence can occur between and within people (Swain, 1964). In this study, dependence between and within people and between people-machines or between events-other events is extended to the same principle. The dependence means how the probability of failure or success of one task can be related to the failure or success of another task. If the conditional probability of one event is the same regardless of whether another event occurs, the two events are independent, otherwise dependent. Conditional probabilities may be applied differently depending on the degree of dependence. However, this paper assumes that the relationship between the two components or events is independent to simplify the calculation process. Once the configuration and dependence of the system components are defined, the formulas in ***Table 9. 3*** are applied to calculate the reliability of the sub-system and the entire system.

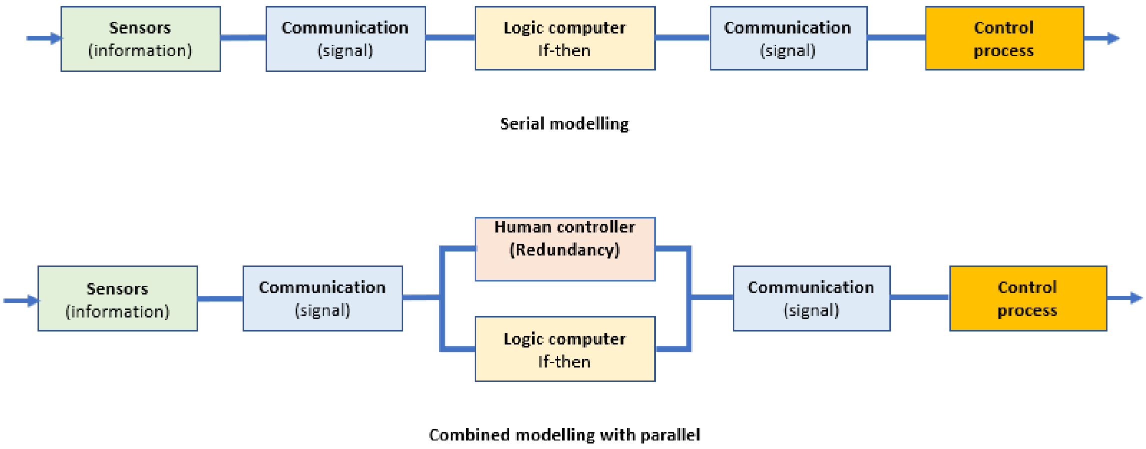

Figure 9. 3 Example of serial and parallel system modelling

Table 9. 3 System reliability corresponding system configuration and dependency

System description	**System sub-task dependency**	**Notation for Error Probability & Reliability**	
Parallel system	Dependency	$EP_{Task} = Min\{EP_{Sub-task\ i}\}$ or	(5)
		$R_{Task} = Max\{R_{Sub-task\ i}\}$	(6)
	Independency	$EP_{Task} = \prod(EP_{Sub-task\ i})$ or	(7)
		$R_{Task} = 1 - \prod(1 - R_{Sub-task\ i})$	(8)
Series system	Dependency	$EP_{Task} = Max\{EP_{Sub-task\ i}\}$ or	(9)
		$R_{Task} = Min\{R_{Sub-task\ i}\}$	(10)
	Independency	$EP_{Task} = 1 - \prod(1 - EP_{Sub-task\ i})$	(11)
		$\approx \sum(EP_{Sub-task\ i})$ or	(12)
		$R_{Task} = \prod(R_{Sub-task\ i})$	

9.4 Application of hybrid method combined the STPA and the SLIM to ESD system of the LNG ship-to-ship bunkering process (Case study 4)

9.4.1 LNG bunkering process overview

On January 1, 2020, new regulations governing Sulphur emission limitations from ships became effective, following the MARPOL Annex IV amendment (IMO, 2018a). The primary change of the MARPOL Annex VI is the addition of emission control zones (ECA) to gradually reduce Sulphur Oxides (SOx), Nitrogen Oxides (NOx), and Particulate Matter (PM) emissions globally and to reduce air pollutant emissions in designated waters to improve global air quality, preserve the environment, and protect human health. Several potential alternatives to traditional marine fuels, such as abatement technologies and alternative marine fuels, have been introduced to the maritime industry over the previous two decades (Jang et al., 2021). In this context, LNG is being accepted as an alternative fuel for ships as a strategy for environmental compliance for vessels during navigation and port operations. LNG as a ship fuel has an immediate and significant impact on reducing SOx, PM, and NOx emissions. As a result, the applicable multilayer regulatory framework strongly favours the usage of LNG as fuel (EMSA, 2018). In addition, global initiatives to safeguard the environment will enhance the trend toward LNG-powered fleets and the need for LNG bunkering at the port.

In contrast, the rising concern is that if LNG becomes widely employed as a ship fuel, the degree of risk associated with bunkering and the general procedures used in containment and operation would considerably increase. LNG is well recognised as a clean fuel that can be consumed entirely and effectively, with very little soot produced during small-scale combustion (Sun et al., 2014). On the other hand, LNG vapour in the air is explosive under certain concentration limits. Once ignited, free natural gas clouds burn very slowly, resulting in comparatively little overpressure in open space. However, if flammable natural gas is generated in a confined space, the surrounding areas may experience higher overpressure.

Fires and explosions are the primary dangers associated with LNG storage and bunkering, and they may occur due to leaks and spills in the presence of ignition sources (Aneziris et al., 2020).

Furthermore, liquefied natural gas is a cryogenic liquid stored at a very low temperature of -162°C at atmospheric pressure. The cryogenic liquid that comes into touch with the hull structure will cause the fragile hull to fracture and lose ductility, destroying the ship's structure (Li and Huang, 2012). Moreover, it can cause cryogenic burns to human skin and an asphyxiant in an enclosed space. Therefore, LNG should be handled by establishing a very high level of safety measures and robust procedures. However, concerns about this risk are due to the complex system of LNG fuel ships and the feature that LNG bunkering progress in the interaction of several stakeholders with a different contexts. In system reliability assessment, humans are an inevitable component to consider since they play a significant part in increasing safety onboard; reliability assessment has always been a critical subject for researchers and decision-makers in this field (Kayisoglu et al., 2021).

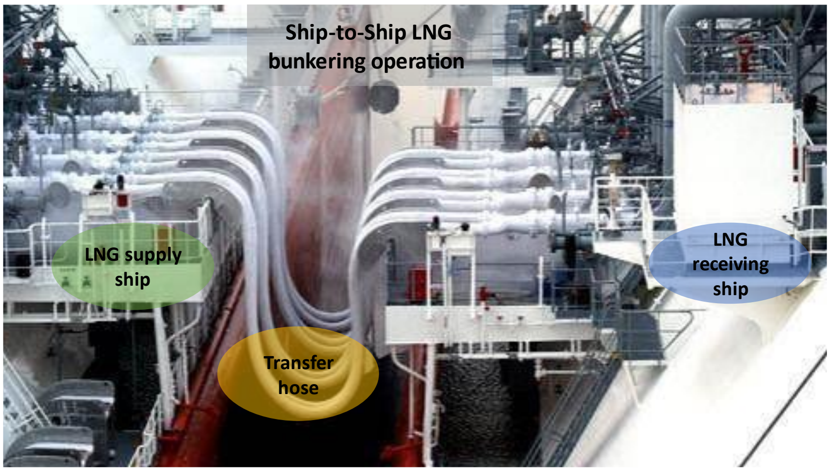

Figure 9. 4 Ship-to-Ship LNG bunkering operation(KLAW, 2022)

Ship-to-ship LNG bunkering is supplying LNG from a bunker supply ship to a bunker receiving ship that uses LNG as a propulsion fuel via a transfer hose, as illustrated in ***Figure 9. 4***. The bunkering procedure may be separated into three phases: the pre-bunkering phase, which involves safe mooring and hose connection, the bunkering phase, which consists in filling LNG, and the post-bunkering phase. While bunkering methods vary according to ship and facility, the following general sequences apply (EMSA, 2018).

Step1. Initial Precooling

Before starting the operation, precooling the filling lines and the cargo pump at the discharging unit is necessary.

Step 2. Connection of Bunker Hose

After the previous precooling is complete, the transfer hoses are attached to the manifold. Sophisticated hose handling equipment like hose cranes or loading arms may be used to convey bunker hoses to the receiving vessel. Each manifold must be earthed, and an insulating flange near the coupling must be put on the receiving vessel to avoid ignition sources caused by electrostatic build-up.

Step 3. Inerting the Connected System

The inerting procedure involves injecting an inert gas into a system to substitute a hazardous gas already present. Nitrogen is used as an inerting gas to eliminate moisture and oxygen from storage tanks and the connecting pipe. In particular, the presence of oxygen in the system causes an explosive environment within the LNG supply line, resulting in potentially hazardous scenarios that should be prevented using an inerting process.

Step 4. Purging the Connected System

For the remaining nitrogen to be removed from the system following engine specifications, the system is purged with natural gas until the ratio is between 97–98 per cent.

Step 5. LNG Filling

When all the necessary preparations have been completed, the LNG filling process may begin. There are two different methods of bottom filling and top filling in the filling sequence.

Step 6. Liquid Line Stripping

After the pump has been turned off, the liquid collected in the bunker hoses must be discharged before the disconnection can be made.

Step 7 Inerting

In a process similar to Step 3, the LNG bunkering line should be inerted before disconnection at the end of the operation.

The strategy and planning for emergency events that may happen throughout the LNG bunkering operation are critical for protecting workers, the environment, the public, and assets in the case of an accident. Thus, building and executing appropriate LNG systems and bunkering operations (ABS, 2017).

The Emergency Shutdown (ESD) system is critical to the vessel's safety. The ESD system is installed as part of an LNG bunkering system designed safely and effectively to stop the flow of LNG (vapour as applicable) or prevent damage to the delivery system in an emergency. The control systems involved in the ESD, which is a linked system to allow both parties (onboard receiving ship and the bunkering ship) to shut down the transfer in an emergency, can be activated automatically or manually. ESD can be composed of two parts (ABS (2017); IACS (2017); EMSA (2018)). The ESD stage 1 system shuts the LNG transfer process down in a controlled manner when it receives inputs from one or more hazardous events listed in ***Table 9. 4***. While the ESD stage 2 is a system that activates the decoupling of the transfer system between the two vessels. Therefore, risk analysis for the ESD system should also consider the entire ship-to-ship LNG bunkering process. The ESD system is configured as a sub-system, and interactions between humans and machines, including software should also be investigated for analysis.

9.4.2 System description: LNG bunkering and emergency shut down system

The system is to be analysed, and the boundary should be defined to identify system hazards (Leveson and Thomas, 2018). The abstraction of the Emergency Shutdown (ESD) system for ship-to-ship LNG bunkering is conceived in ***Figure 9. 5***, and the interaction between humans and machines is modelled in ***Figure 9. 6*** to support the system definition. The initial stage of the conceptual visual definition has been developed into a more detailed physical diagram in ***Figure 9. 7*** to support a unified perspective and understanding of the system by experts for LNG bunkering operations. LNG Bunkering is the practice of providing Liquefied Natural Gas (LNG) fuel to LNG fuelled ships. Depending on the LNG bunkering mode, it can be divided into ship-to-ship, truck-to-ship, or terminal-to-ship. In this study, the ship-to-ship bunkering mode was selected. A ship that supplies fuel is called a bunker supply vessel, and a ship that consumes LNG as fuel and receives it is called a receiving vessel. When the two ships are safely moored for LNG fuel supply, the two types of hoses are connected to the manifold flanges of both ships. One is a liquid filling hose, and the other is a vapour return hose. An operator is in the control room of each vessel for LNG bunkering operation, and a site operator is also located near the manifold. A control monitoring system that controls the process of the system is installed on each ship. The LNG bunker transfer system is to be equipped with a linked and compatible ESD system that is completely independent of the installed control and monitoring system. This system will be used to halt bunker flow in an emergency (ABS, 2017). ESD systems can be activated automatically and manually on each vessel, from both the control room and manifold side, and ESD systems on both vessels are linked. Emergency release coupling (ERC) is installed at each hose to disconnect the fuel supply in an emergency immediately. Remotely operated ESD valves must be installed in each bunkering line immediately adjacent to the manifold joining point.

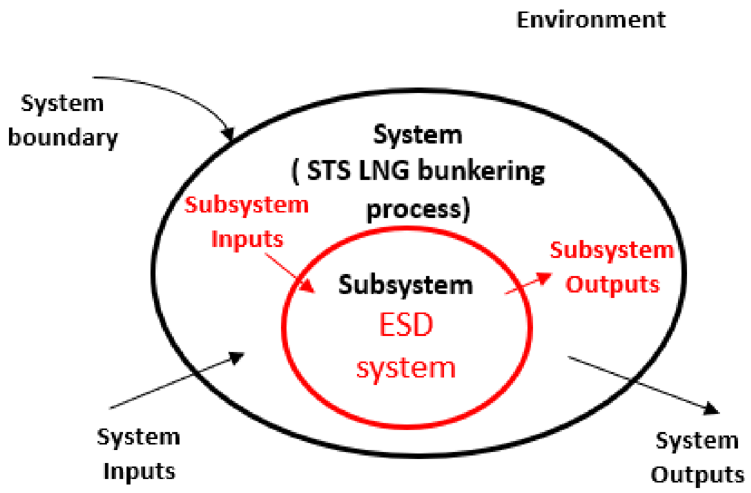

Figure 9. 5 Abstraction of ESD system for ship-to-ship LNG bunkering process (Adopted from Leveson and Thomas (2018))

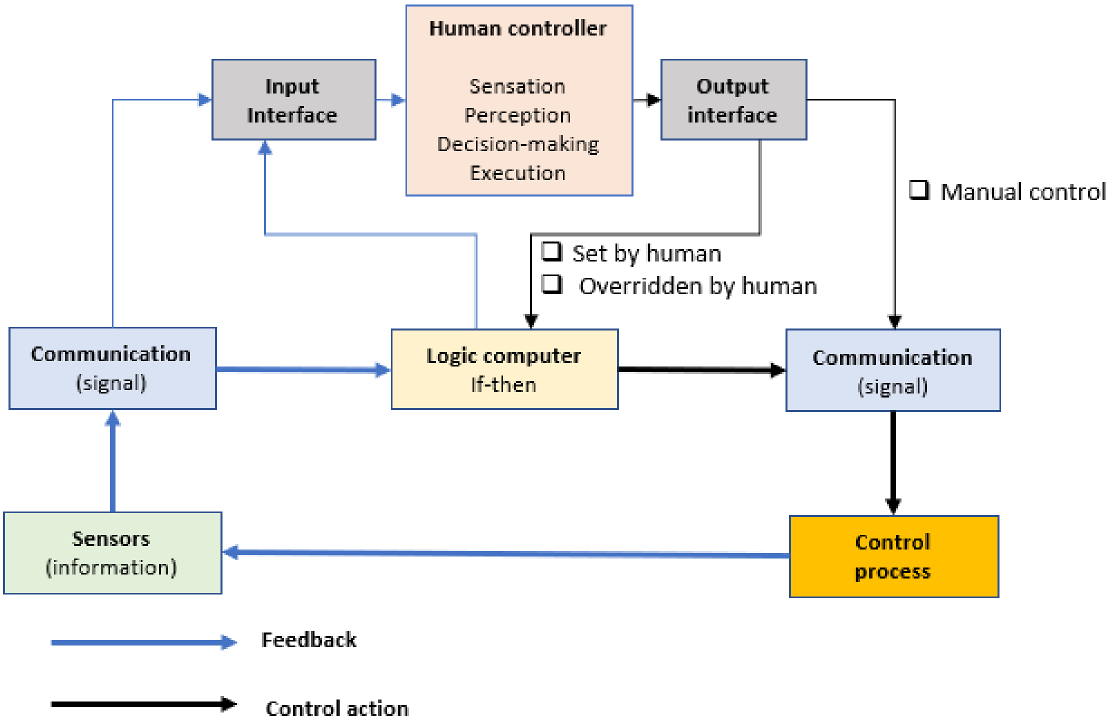

Figure 9. 6 Human-machine interaction model

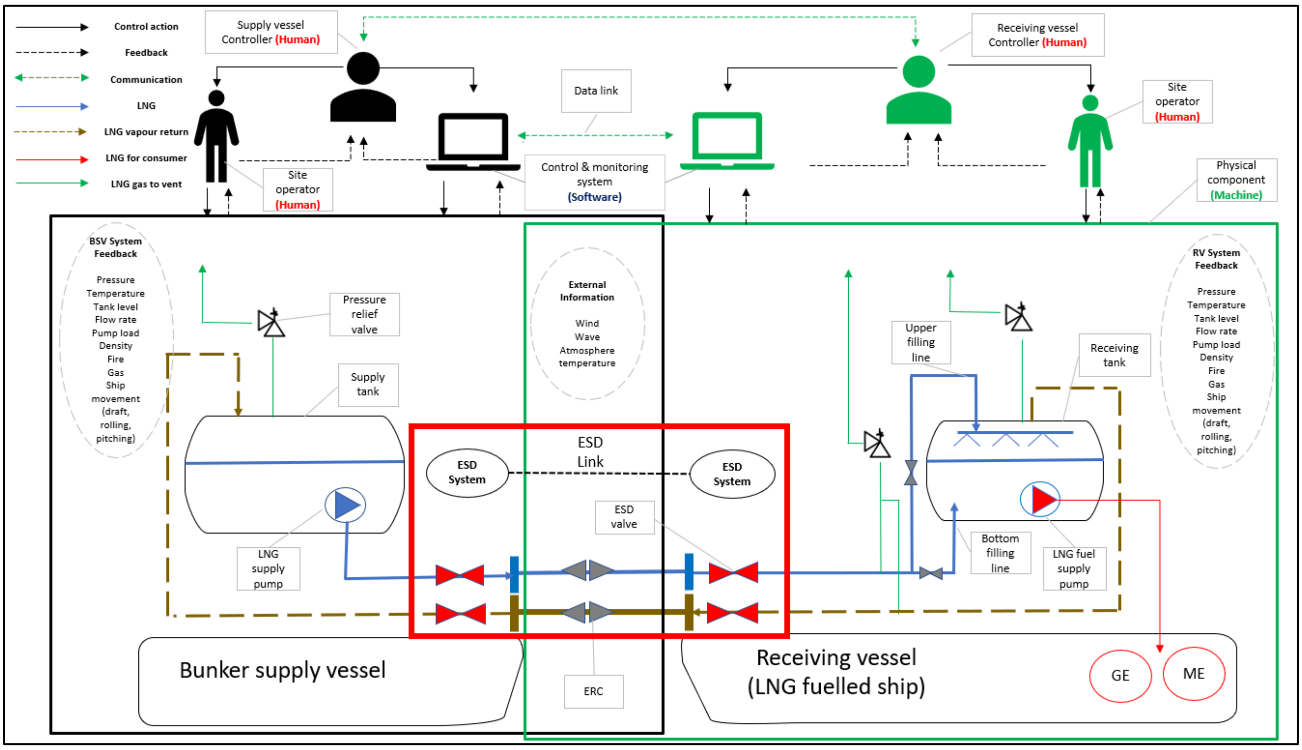

Figure 9. 7 Simplified ship-to-ship LNG bunkering physical diagram

9.4.3 Loss and hazard identification

Losses may include death or injury, ship structural damage, marine pollution, mission failure, or any other type of loss deemed undesirable by stakeholders (Leveson and Thomas, 2018). During ship-to-ship LNG bunkering, loss and hazards of the ESD system are defined as shown in ***Table 9. 4***. Hazards are defined as events that the ESD system that must be activated according to the Guide for LNG bunkering (ABS, 2017), and described as loss as a state in which the ESD system fails to emergency response when such a hazardous situation occurs.

Table 9. 4 Loss and Hazards definition

Definition of System Loss and hazards
Loss
L1. Failure of emergency response by the ESD system when the hazardous situation occurred.
Hazards
H1. Detection of gas in the cargo machinery space at levels more than 60% of LEL
H2. Detection of gas in the bunkering manifold area at levels more than 60% of LEL
H3. High pressure is generated at bunkering manifolds
H4. High-high pressure is generated in the vapour return line
H5. Loss of motive power for the ESD valve
H6. Activation of the emergency release system (ERS) by default
H7. The liquid level in the LNG receiving tank has risen to a High-high level.
H8. High-high pressure is generated in LNG receiving tank
H9. ESD signal generated manually or automatically by LNG receiving vessel
H10. Fire detection onboard

9.4.4 Model control structure

The control structure is a system model composed of feedback control loops that impose safety restrictions on the system's behaviour. A process model determines which control actions are required to maintain the system's effectiveness and explain and anticipate interactions between humans, their mental models, and the logic control system (Leveson and Thomas, 2018). As shown in ***Figure 9. 8***, the control structure for the ESD system represents the responsibilities of humans and software according to the process model. The corresponding control action can be identified from the control structure for the ESD system.

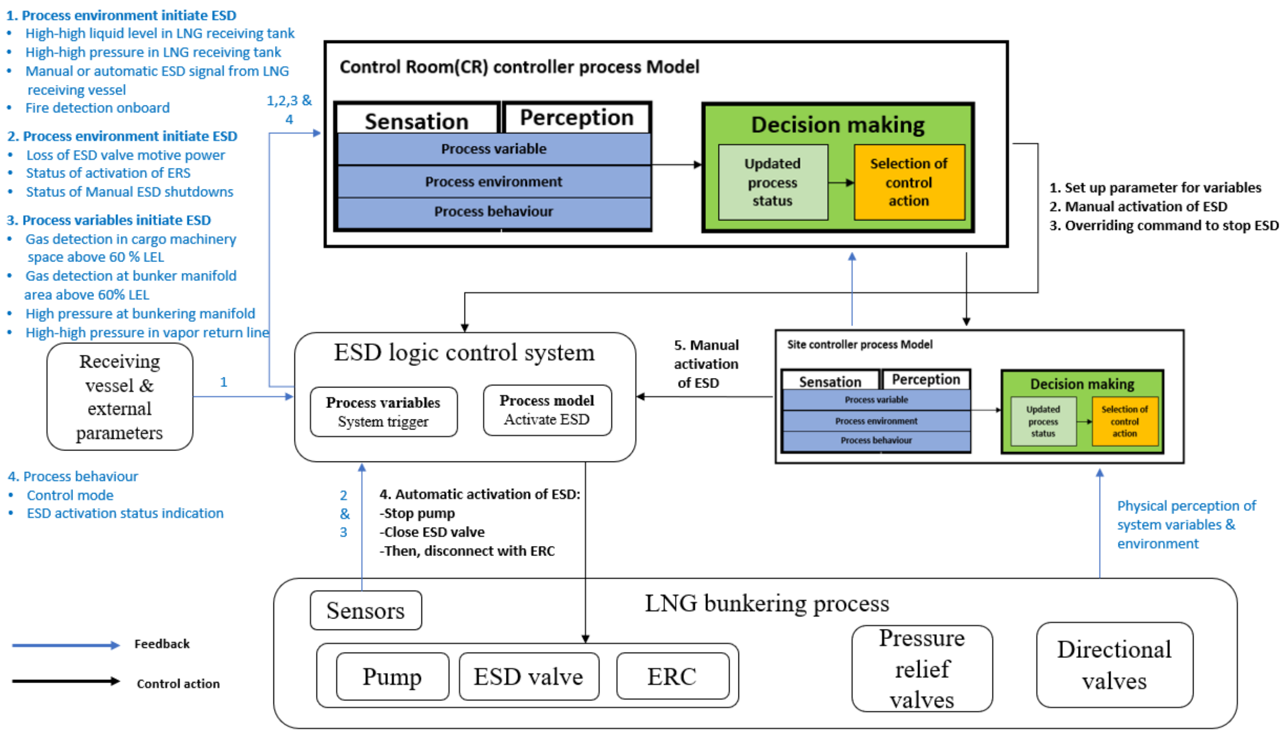

Figure 9. 8 Control diagram for ESD system

9.4.5 Identification of the unsafe control actions

After modelling the control structure, the next phase is identifying Unsafe Control Actions. The system is managed automatically or manually according to the system logic by two human controllers, a control room operator, a site operator near the manifold, and a software controller. This stage identifies Unsafe Control Activities, which are control actions that could result in a hazard in a particular context and worst-case setting. As illustrated in ***Table 9. 5***, hazardous control actions are recognised by possible error modes, including not providing controls, incorrect controls, and timing errors.

Table 9. 5 Unsafe Control Actions (Part of, Full UCAs are listed in Appendix B)

Control action (Controller)	Not providing Cause hazard	Providing Causes Hazard	Timing error Too early/ late
CA1. Setting up user-configurable parameters in the system (Control room operator)	UCA1.1 Operator did not set up system user-configurable parameter for max gas detection level in cargo machinery space. As a result, when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, ESD was not automatically activated by the ESD system. (H1)	UCA1.2 Operator wrongly set up system user-configurable parameter for max gas detection level in cargo machinery space. As a result, when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, ESD was not automatically activated by the ESD system. (H1)	N/A
	UCA1.3 Operator did not set up system user-configurable parameter for max gas detection level at bunkering manifold area. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60% LEL, ESD was not automatically activated by the ESD system. (H2)	UCA1.4 Operator wrongly set up system user-configurable parameter for max gas detection level at bunkering manifold area. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60% LEL, ESD was not automatically activated by the ESD system. (H2)	N/A
CA.2 Activate ESD manually(Control room operator)	UCA2.1 Operator did not activate ESD manually when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, and other controllers did not activate ESD. (H1)	UCA2.2 Operator did activate ESD manually when the amount of LNG vapour gas contained in cargo machinery was within the acceptable limit. (H1)	UCA2.3 Operator did activate ESD manually too late when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, and other controllers did not activate ESD. (H1)

	UCA2.4 Operator did not activate ESD manually. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60%, LEL and ESD were not activated by other controllers. (H2)	UCA2.5 Operator did activate ESD manually when the amount of LNG vapour gas contained near the bunkering manifold was within the acceptable limit. (H2)	UCA2.6 Operator did activate ESD manually too late. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60%, LEL and ESD were not activated by other controllers. (H2)
CA.3 Override to stop ESD activation(Control room operator)	N/A	UCA3.1 Operator did override to stop ESD activation by the system when the LNG vapour gas in the cargo machinery space exceeded 60% LEL. (H1)	N/A
	N/A	UCA3.2 Operator did override to stop ESD activation by the system. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60% LEL. (H2)	N/A
CA.4 Activate ESD automatically(ESD logic computer)	UCA4.1 ESD logic controller did not activate ESD when the LNG vapour gas in the cargo machinery space exceeded 60% LEL. (H1)	UCA4.2 ESD logic controller activated ESD when the amount of LNG vapour gas contained in cargo machinery was within the acceptable limit. (H1)	N/A
	UCA4.3 ESD logic controller did not activate ESD. As a result, when the LNG vapour gas near the bunkering manifold exceeded 60% LEL. (H2)	UCA4.4 ESD logic controller activated ESD when the amount of LNG vapour gas contained near the bunkering manifold was within the acceptable limit. (H2)	N/A

CA.5 Activate ESD manually(Site operator)	UCA5.1 Site operator did not activate ESD manually when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, and other controllers did not activate ESD. (H1)	UCA5.2 Site operator activated ESD manually when the amount of LNG vapour gas contained in cargo machinery was within the acceptable limit. (H1)	UCA5.3 Site operator did activate ESD manually too late when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, and other controllers did not activate ESD. (H1)
	UCA5.4 Site operator did not activate ESD manually when the LNG vapour gas near the bunkering manifold exceeded 60% LEL, and other controllers did not activate ESD. (H2)	UCA5.5 Site operator activated ESD manually when the amount of LNG vapour gas contained near the bunkering manifold was within the acceptable limit. (H2)	UCA5.6 Site operator did activate ESD manually too late when the LNG vapour gas near the bunkering manifold exceeded 60% LEL, and other controllers did not activate ESD. (H2)

9.4.6 Identify loss scenarios

Loss scenarios can be considered the catalysts for hazardous control behaviour, i.e., scenarios that result in UCA and scenarios in which control actions are done wrongly or not done. The causes of unsafe control behaviour are divided into controller failure and decision-making error. The scenarios in which control actions are incorrectly executed (or are not executed at all) result from control path and control process issues. ***Figure 9. 9*** shows multi-level hazards with pathways to cause hazardous scenarios. The human controller model was used to describe unsafe control behaviour in more detail. However, the process of identifying loss scenarios is a highly iterative task. Since the proposed framework for system reliability does not require complete STPA analysis, this section briefly describes how human models and PSFs in ***Table 9. 1*** are utilised to create loss scenarios. The causal scenarios that result in unsafe control actions are described in ***Table 9. 6***.

Table 9. 6 Causal Scenarios result in unsafe control actions

Unsafe Control Actions	Human diagnosis process	Type of Information	PSFs	Example scenarios (Written Form)
UCA2.1	Sensation	Process variables	Input interface (Sound alarm)	The operator did not activate ESD manually because the CR operator failed to hear alarm sounds due to lower sound volume in the control room when the LNG vapour gas in the cargo machinery space exceeded 60% LEL.
UCA2.1	Sensation	Process variables	Working environment (Noise)	The operator did not activate ESD manually because the CR operator failed to hear alarm sounds due to the noise in the control room when the LNG vapour gas in the cargo machinery space exceeded 60% LEL.
UCA2.1	Perception	Process variables	Input interface (Sound alarm)	The operator did not activate ESD manually because the alarm sound was not distinguished when the LNG vapour gas in the cargo machinery space exceeded 60% LEL, so the CR operator was unaware of which alarm sound was activated.
UCA2.1	Perception	Process behaviour	Input interface (Visual indication of control mode)	The operator did not activate ESD manually because the indication of control mode was not distinguished when the LNG vapour gas in the cargo machinery space exceeded 60% LEL. Hence, the CR operator believed that the ESD system was activated automatically, even though ESD failed to be activated.

UCA3.5	Decision-making	Process environment	Procedure (Emergency procedures)	The operator wrongly overrode to stop ESD activation when the action should not have been done because the CR operator believed that ESD activation was not required to be activated when ESD valve motive power was lost due to the wrong procedures.

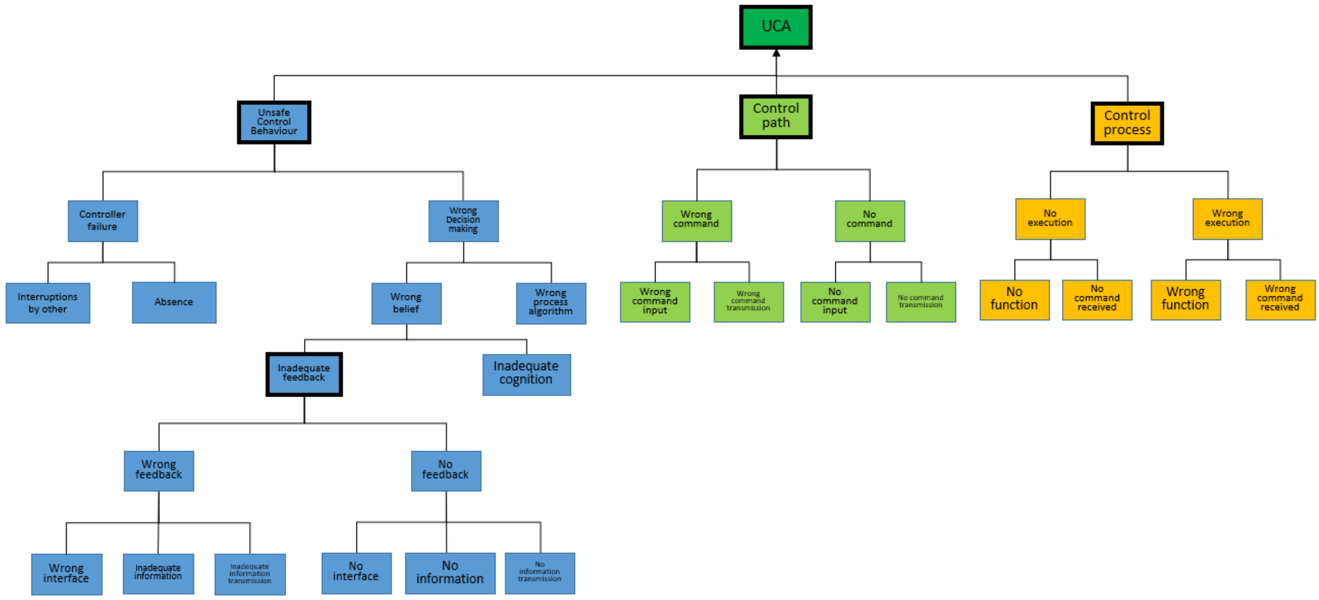

Figure 9. 9 Loss scenario pathway

9.4.7 Human error calculation by the SLIM method

The SLIM was used to quantify human error probabilities for human responsibilities for incorporating them into the probabilistic risk assessment. Five professionals with practical experience in ship-to-ship LNG bunkering as crew members or safety system auditors did this assessment. Following that, each expert conducted the review separately to eliminate groupthink. The standard procedures (ABS (2017); EMSA (2018); IACS (2017)) and findings through STPA analysis were provided for evaluation. The experts were tasked with responding to questionnaires designed to ascertain the relative importance and evaluation of corresponding PSFs.

9.4.7.1 Identification of human tasks

Through the control structure of STPA analysis, the tasks of the control room operator and site operator are identified in ***Table 9. 7***. First, it is necessary to analyse what context human role is required for more careful human error prediction. This is divided before and after the start of the LNG process. Before the process, the human role is to set up system parameter values for events that lead to hazards in the ESD system process. Next, the required human tasks during the process show what actions are needed in which system context and who is responsible, as shown in ***Table 9. 8***.

Table 9. 7 List of human tasks for ESD system during LNG ship-to-ship bunkering operation

Tasks
T1. Set*ting up user configurable parameters in the system* by CR operator
T2. Activate ESD manually by the CR operator
T3. Overring to stop ESD by the CR operator
T4. Activate ESD manually by the site operator

Table 9. 8 human responsibility per system context for ESD system

System variables/ environment	Automatic system behaviour	Expected action	Responsible person
Before start	N/A	Set up parameters	CR controller
Normal	Activated	Overriding	CR controller
Normal	Not activated	No action required	N/A
Abnormal	Activated	No action required	N/A
Abnormal	Not activated	Activate ESD	CR controller and site controller

9.4.7.2 PSF weighting

At this stage, the priority of PSFs was applied to consider the impact of each PSF on human performance. Again, linear scales from 0 to 100 were used for evaluation, and the determined mean weights were normalised to indicate relative importance. Since the degree of PSFs contributing to each task is different, the relative importance of PSFs for each task was evaluated separately, as shown in ***Table 9. 9***.

Table 9. 9 Weightings of PSFs per task

Task	PSF Weighting		EX1	EX2	EX3	EX4	EX5	Mean Weight	Normalised Weight(W_i)
Task1	PSF1	Interface (Input device)	10	20	10	10	10	12	0.04
	PSF2	Interface (Output device)	80	70	80	70	90	78	0.23
	PSF3	Procedure	50	60	70	60	60	60	0.18
	PSF4	Working condition	30	50	70	40	30	44	0.13
	PSF5	Time	10	20	10	10	10	12	0.04
	PSF6	Experience & training	70	60	80	60	70	68	0.20
	PSF7	Environmental condition	10	10	10	10	10	10	0.03
	PSF8	Complexity	10	20	30	20	10	18	0.05
	PSF9	Organisational factors	20	30	40	20	40	30	0.09
Task2	PSF1	Interface (Input device)	80	80	90	90	80	84	0.20
	PSF2	Interface (Output device)	90	80	90	70	90	84	0.20
	PSF3	Procedure	70	60	70	60	70	66	0.16
	PSF4	Working condition	50	60	50	40	60	52	0.12
	PSF5	Time	30	40	50	40	40	40	0.09
	PSF6	Experience & training	50	60	50	40	40	48	0.11
	PSF7	Environmental condition	10	10	10	10	10	10	0.02
	PSF8	Complexity	10	10	30	40	20	22	0.05

	PSF9	Organisational factors	10	10	20	10	30	16	0.04
Task3	PSF1	Interface (Input device)	80	80	90	90	80	84	0.19
	PSF2	Interface (Output device)	90	80	90	70	90	84	0.19
	PSF3	Procedure	70	60	70	60	70	66	0.15
	PSF4	Working condition	50	60	50	40	60	52	0.12
	PSF5	Time	30	40	50	40	40	40	0.09
	PSF6	Experience & training	70	70	80	80	60	72	0.16
	PSF7	Environmental condition	10	10	10	10	10	10	0.02
	PSF8	Complexity	10	10	30	40	20	22	0.05
	PSF9	Organisational factors	10	10	20	10	30	16	0.04
Task4	PSF1	Interface (Input device)	70	80	80	90	80	80	0.15
	PSF2	Interface (Output device)	80	80	90	90	70	82	0.15
	PSF3	Procedure	60	70	80	90	80	76	0.14
	PSF4	Working condition	70	80	70	80	90	78	0.14
	PSF5	Time	30	40	50	30	50	40	0.07
	PSF6	Experience & training	60	70	80	70	70	70	0.13
	PSF7	Environmental condition	50	60	90	80	70	70	0.13
	PSF8	Complexity	10	30	50	40	20	30	0.06
	PSF9	Organisational factors	10	20	10	10	10	12	0.02

9.4.7.3 PSF rating and SLI

The PSF rating process is crucial in calculating human errors, but the PSF rating is ambiguous, and the gap is wide depending on experts. For this reason, a Likert scale was provided, as mentioned in section 3.2.2. The selected expert group has the relative importance (w_j) which was assigned to 0.20, 0.18, 0.21, 0.20 and 0.21, according to evaluation criteria such as professional position, serviced time, and experience. Consensus rating R_i for PSF_i is computed for each task using equation (9-2), as demonstrated in ***Table 9. 10***, which shows values of R_1 to R_9 for task 1 based on expert judgments on PSF_i. After the consensus rating

R_i and normalised weight W_i for PSF_i are determined, the Success likelihood Index (SLI) for all tasks was derived by equation (9-3), as shown in

Table 9. 11.

Table 9. 10 Experts' PSFs ratings and consensus ratings for Task1

Expert No.	**Ex1**	**Ex2**	**Ex3**	**Ex4**	**Ex5**	**Consensus**
$\boldsymbol{w_j}$	0.2	0.18	0.21	0.2	0.21	**Rating** $\boldsymbol{R_i}$
PSF1	78	86	78	73	85	80
PSF2	50	71	76	54	68	64
PSF3	75	50	70	73	50	64
PSF4	70	74	75	50	73	68
PSF5	45	50	65	60	73	59
PSF6	70	80	65	53	75	68
PSF7	50	52	70	80	65	64
PSF8	50	30	45	30	50	41
PSF9	50	75	70	50	50	59

Table 9. 11 SLI for each task

Task		**PSF1**	**PSF2**	**PSF3**	**PSF4**	**PSF5**	**PSF6**	**PSF7**	**PSF8**	**PSF9**	**SLI**
Task1	R_i	80	64	64	68	59	68	64	41	59	64.08
	W_i	0.04	0.23	0.18	0.13	0.04	0.2	0.03	0.05	0.09	
Task2	R_i	68	80	59	68	41	68	64	36	59	64.50
	W_i	0.2	0.2	0.16	0.12	0.09	0.11	0.02	0.05	0.04	
Task3	R_i	68	80	54	68	41	68	64	36	59	63.97
	W_i	0.19	0.19	0.15	0.12	0.09	0.16	0.02	0.05	0.04	
Task4	R_i	36	64	54	31	41	54	41	24	59	45.37
	W_i	0.15	0.15	0.14	0.14	0.07	0.13	0.13	0.06	0.02	

9.4.7.4 Human error calculation

The human error probability is derived from SLI by calculating anchor values and performing the calibration equation (9-4). In the case of LNG ship-to-ship bunkering work, absolute probability judgment by experts was used for endpoints because there is no empirical data available for human failure probabilities. This method is used in rare event scenarios to estimate calibration tasks (Kayisoglu et al., 2021). This method allows experts to assume the best and worst scenarios to estimate constants 'a' and 'b'. For the control room operator and site operator, the constant values were determined as shown in ***Table 9. 12*** by experts considering the given context. Then, human error probability is derived by equation (9-4), as shown in ***Table 9. 13***.

Table 9. 12 Estimate HEP for the best and worst scenarios of LNG ship-to-ship bunkering operation

Responsible person	HEP Best scenario	HEP Worst scenario	Constant a value	Constant b value
Control room operator	1.00E-04	1.00E-02	4.32E-05	-4.36E-03
Site operator	1.00E-03	1.00E-01	4.53E-04	-4.58E-02

Table 9. 13 Human error probability for each task

Tasks	**SLI**	**Log (Success probability)**	**Success probability**	**HEP**
Task1	64.0780	-0.0016	0.9963	3.67E-03
Task2	64.5028	-0.0016	0.9964	3.63E-03
Task3	63.9707	-0.0016	0.9963	3.68E-03
Task4	45.3666	-0.0252	0.9436	5.64E-02

9.4.8 System reliability assessment for human-machine interactive controller

To present human error in the system reliability model, it is necessary to understand how the interactive relationship between humans and machines is connected to the system process. The control roof of the human-machine interaction system for the ESD system for the ship-to-ship LNG bunkering process is defined in ***Figure 9. 8***. The failure probabilities of software, sensors and equipment vary on factors such as built specification, manufacturer, and time dependency of the life cycle. However, considering the limited scope of this study, quantitative data collection for all parts of equipment needs to be assumed reasonably based on literature (Khalaquzzaman et al. (2014); Kang et al. (2009); Kamyab et al. (2013); ABS (2017); EMSA (2018); IACS (2017)). For this analysis, the software's failure probability was assumed to be 0.001. Each sensor, cable, and equipment was supposed to be the same value, but 1.0E-6 for the rest of the equipment except software, reflecting the independent redundancy requirements of the currently applied Maritime Rules. The remained problem is how to interpret each role of human being in a reliability model. First, the failure of the control room operator to override and set system parameters is interpreted in terms of human errors because it degrades the reliability of the software. If any of them fails, the function of the human-software interactive controller fails, so three elements are connected in series.

On the other hand, if the main functional human-software controller fails, the ESD system can be manually activated by the control room operator and the site operator, respectively, so the human role here serves as a redundancy. Therefore, the role of humans in manually activating ESD is connected in parallel. The setup task is linked with the logical computer in serial connection, and their relation is dependent, but the logical computer and the override functions are independent. Considering these three factors, the calibrated error probability 7.34E-03 obtained by formulas in ***Table 9. 3*** becomes the error probability of the human-software interactive controller that carries out the primary automatic activation. The human roles of the control room operator and site operator for manual activation of ESD and main software functions are independent. The total system reliability calculated by the equation in ***Table 9. 3*** is 1.5E-06, considering the system configuration and dependence relationship as shown in ***Figure 9. 10***. This is a relationship in which five components from sensors to equipment are connected in series and dependent, so in the end, the errors of the entire system are represented by maximum errors. In terms of reliability, the minimum reliability represents the overall system reliability. This means that the overall reliability increases only when the error probability of the controller composed of humans and software is decreased to 1.0E-06, which is the error level of other components.

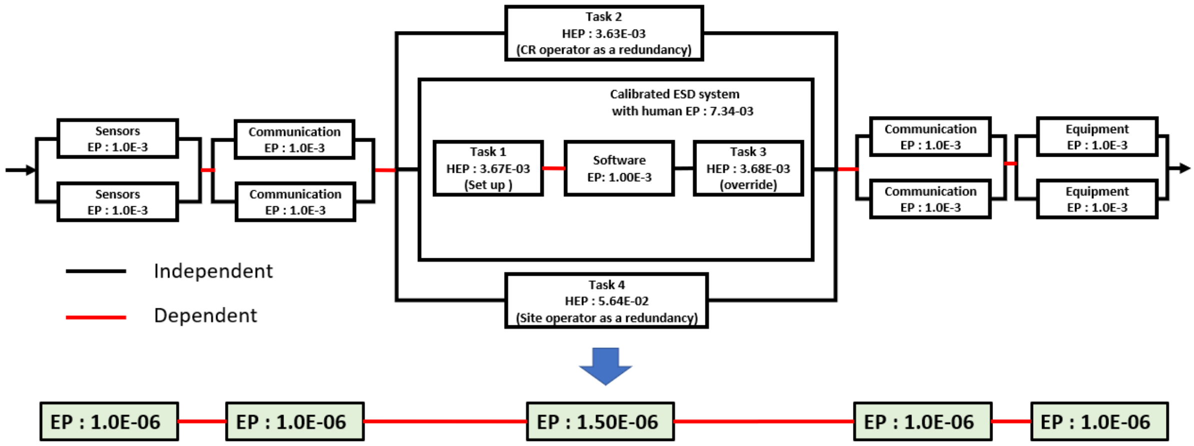

Figure 9. 10 System reliability for ESD system for ship-to-ship LNG bunkering

9.4.9 Comparative analysis for system design alternatives

The current ESD system with an error probability of 1.5E-06 was used as the baseline to improve the controller reliability in which humans and software link. The following three different system configurations were used as alternatives for comparison. The first alternative is to install an additional independent ESD system. The second alternative is to place one supervisor in the control room. Finally, the third alternative is to deploy one supervisor near the manifold site. The system reliability block diagram for each case is illustrated as shown in ***Figure 9. 11***, ***Figure 9. 12***, ***Figure 9. 13*** and ***Figure 9. 14***, and the human-machine controller reliability of each system is obtained as shown in ***Figure 9. 15***. Alternative 2 offers the highest reliability.

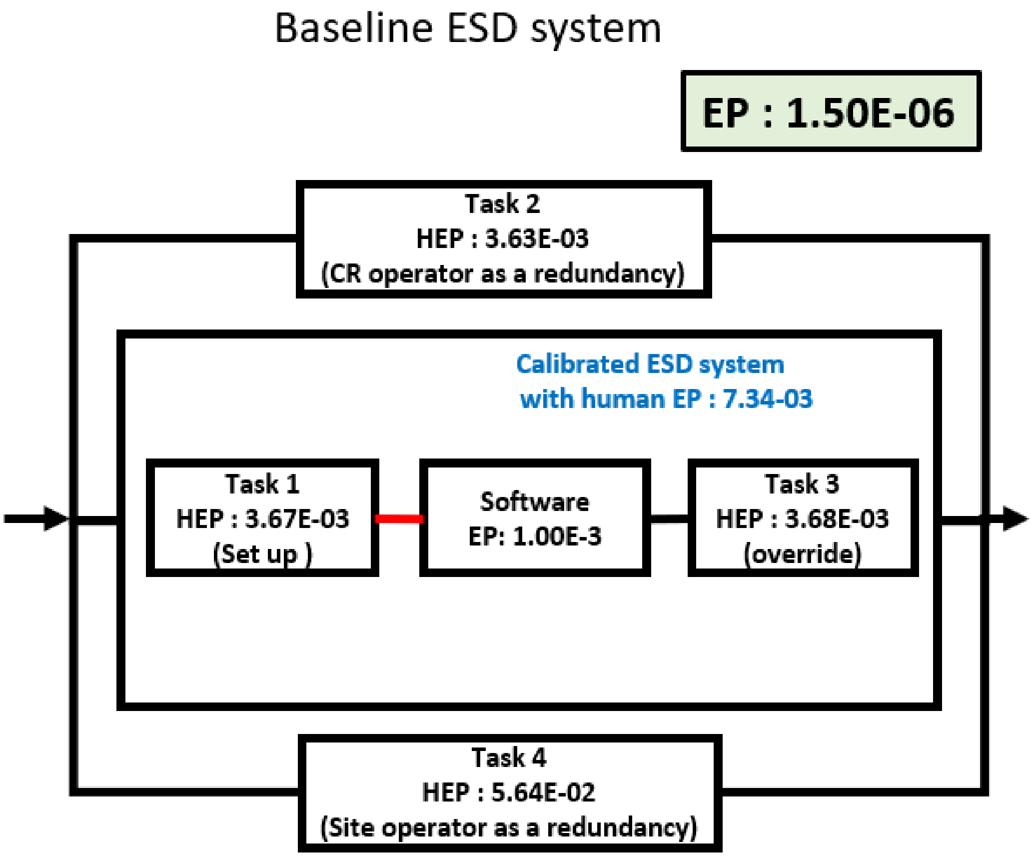

Figure 9. 11 Baseline ESD system reliability block diagram

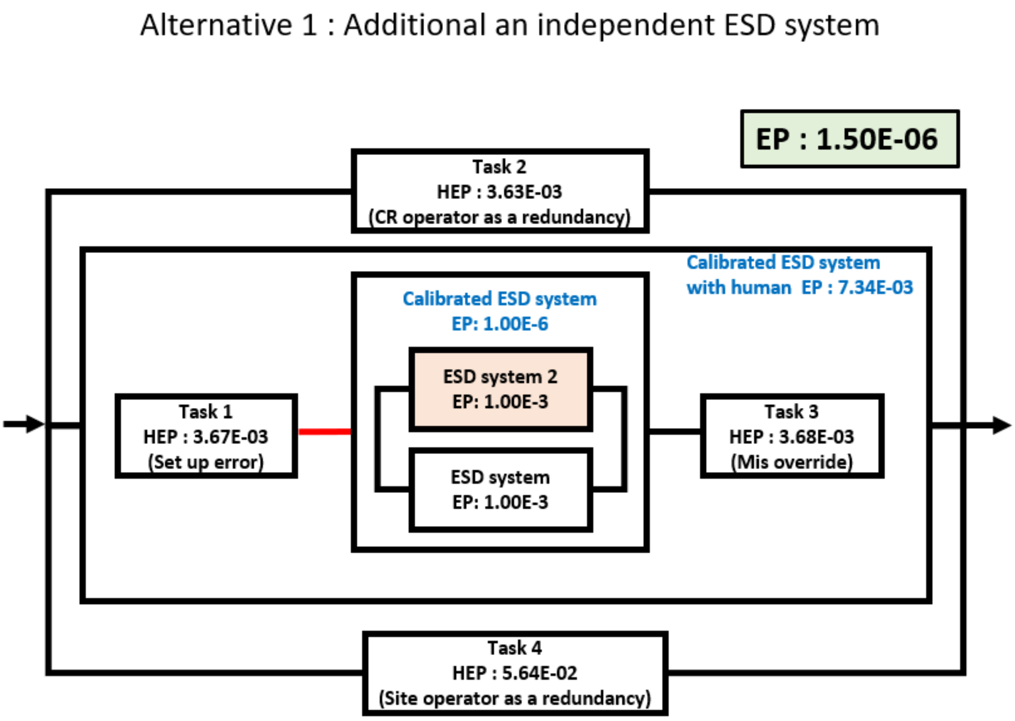

Figure 9. 12 ESD system reliability block diagram (Alternative 1)

Alternative 2
Additional a supervisor in control room

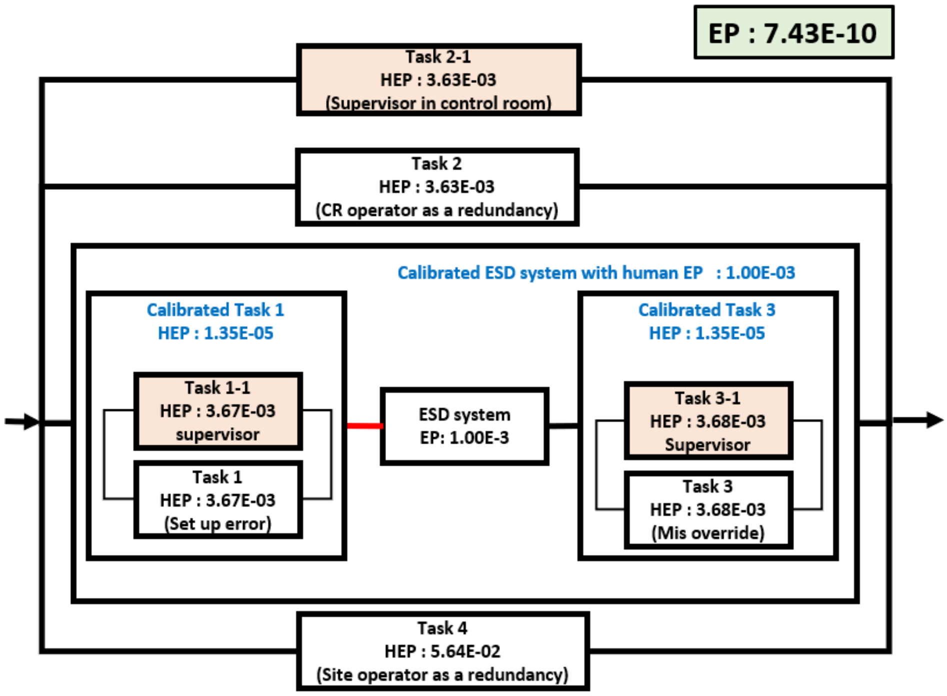

Figure 9. 13 ESD system reliability block diagram (Alternative 2)

Alternative 3
Additional a supervisor at site

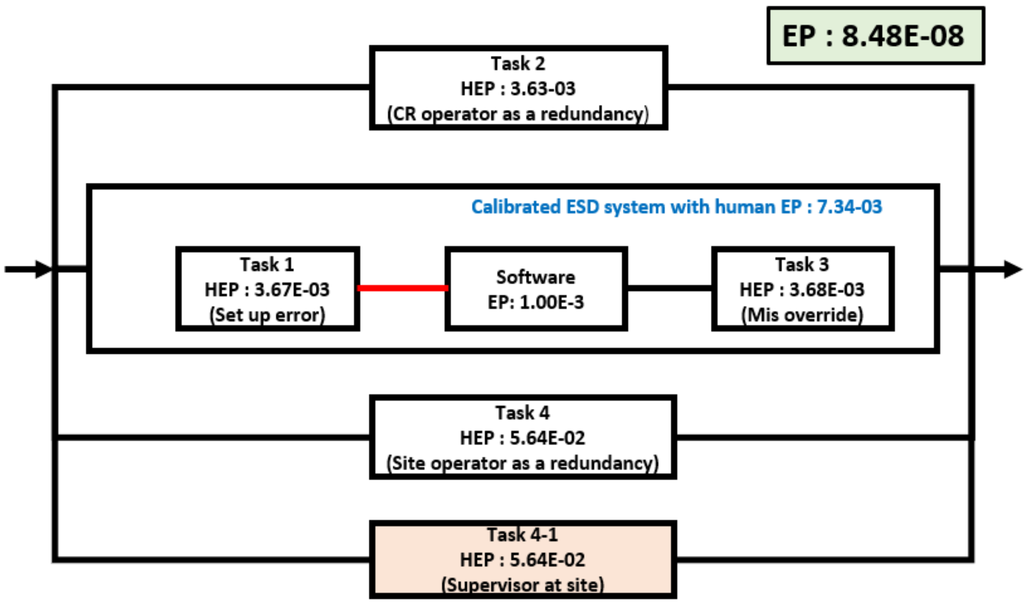

Figure 9. 14 ESD system reliability block diagram (Alternative 3)

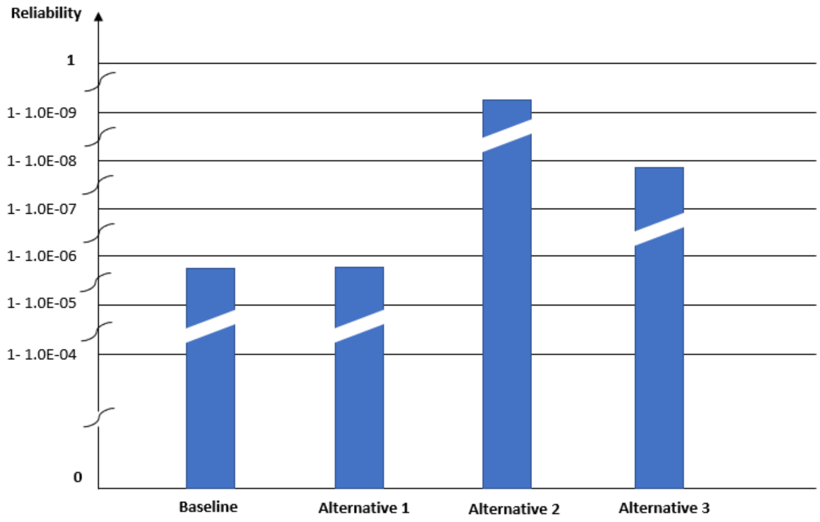

Figure 9. 15 Reliability of the ESD system controllers with different system configurations

9.5 Findings and discussion

The proposed method identified the flawed interaction between humans and other system components through a systematic approach, thereby recognising unsafe control actions that lead to hazardous situations from a system control perspective. The scenarios were explored based on identified UCAs as to what causes contribute and how unwanted dangerous situations occur. In addition, identified hazards, unsafe control actions, and loss scenarios induce safety recommendations for each step. This qualitative analysis contributes to strengthening safety measures through better understanding and interpretation of the system while identifying human responsibilities in the system. According to the system situation, human responsibility identified through system analysis was expressed as expected human roles, as shown in ***Table 9. 8***. This means that within a complex system, the human role is not evaluated independently regardless of the system's situation but should be treated as a dependent human role in response to changes in the system situation. Next, the error probabilities of the identified system components were predicted, and human errors were obtained through the SLIM method. In the process of quantifying human errors, the Likert scale and weighted normalised mean value were used to minimise uncertainty, ambiguity, and inconsistency according to expert groups, and experts carefully evaluate human error probability. Human errors from task 1 to task 4 were evaluated from a minimum of 3.68E-03 to a maximum of 5.64E-02. All the tasks 1 to 3 performed by control room workers showed similar probabilities of error at 3.6E-03.

In contrast, the probability of human error at the site operator was 5.64E-02, indicating that the overall hostile working environment, shown by experts' PSF evaluation, had a more significant impact on the site operators' performance than the control room operators. When the probability of error in individual system components was obtained, they were modelled by the system reliability block diagram to evaluate the reliability of the entire system. Thus, individual system components constitute the system by two criteria, system description, serial or parallel and level of dependence. Finally, a comparative analysis of three different design options was conducted against the reliability of the integrated controller between humans and software analysed in the case study for design suggestions. The current system's reliability was (1 − 1.50E-06), while the reliability was the highest at (1 − 7.43E-10) when one more supervisor was placed in the control room and the reliability level of (1 − 8.48E-08) when one more supervisor was placed on the site. However, even if an independent ESD is installed, the overall reliability does not change. These results show how critical human roles are in a complex system where humans, machines, and software interact. Humans have a more significant impact on the system than any other system element if humans have an overriding authority over the system. This human influence can work positively or negatively on system performance from a safety perspective. Therefore, it is necessary to recognise that the level of each factor affecting human performance plays a decisive role in system reliability. Consequently, it is essential to identify flawed interactions between humans and machines and humans' overriding authority that can be misused. In addition, if the human role in the system is well designed and the factors affecting human performance are well managed, the human role in the system will play a more significant role in recovering system errors than contributing to the cause of the error. These findings can contribute to the prioritisation of system improvement elements while identifying the optimal combination of humans and machines to maximise system reliability.

9.6 Chapter summary

This chapter demonstrated a new hybrid method combining STPA and SLIM to analyse the human role qualitatively and quantitatively in emergency shutdown operations during LNG vessel-to-vessel bunkering. This systemic approach based on STPA was created to assist in understanding human process models and capturing additional causal scenarios. The human process model with PSFs is unique as it proposes a new simplified model of the human diagnosis process from a system perspective. The scenario development process is a newly proposed guideline that can be quickly applied to identify a rich set of scenarios related to human behaviour, including system information, human diagnosis processes, and performance shaping factors. The SLIM calculates quantitative human error probabilities from the identified human responsibilities by measuring the contribution of Performance Shaping Factors to human reliability. Traditional STPA does not pursue an error probability model, but quantification is an inevitable process that should be applied to probabilistic risk assessment frameworks currently used as Maritime's industrial standard. Furthermore, this method can represent errors of all system components to integrate HRA into the whole risk picture through the system reliability block diagram.

In conclusion, in safety-critical systems that involve and rely on human interactions, human reliability assessment alone will not be sufficient to evaluate human behaviour without considering operators' interactions with the system. In such systems, the human operators' role should be viewed as a system component and analysed in relation to other components that interact with it. The approach demonstrated in this chapter show promising results for calculating overall reliability in such operations.

10 Discussions and recommendations for the future research

10.1 Overview

To begin, Section 10.2 of this chapter summarises this research study, demonstrating its distinctiveness through the use of the research findings. Then, section 10.3 contains a complete overview of the aim and objectives of this research. Then, Section 10.4 discusses the shortcomings of this PhD study. Finally, the concluding section, Section 10.5, contains a range of recommendations for future researchers.

10.2 Brief review of the research study and its originality

There is no doubt that human performance is vital in safety-critical industries. Therefore, understanding the mechanism by which humans interact with the system is critical. Numerous academics have made efforts to comprehend human behaviour in terms of reliability. They begin with identifying the components influencing the performance and probabilistically forecasting it. Initially, academics concentrated on quantifying human error and attempting to anticipate it more precisely. Numerous strategies for assessing human reliability have been developed and implemented. The earlier studies focused on characterising human behaviour from an independent human perspective. However, because human performance is influenced by external elements such as the environment and organisation, additional research has been undertaken in this field. As a result of this endeavour, researchers aimed to estimate the likelihood of human error. Apart from the accuracy of the expected human error probability, it is worth noting that the predicted human error probability offers the system very little information. In other words, the likelihood of human error alone is insufficient to account for the system's influence on humans. This is because the success or failure of a human's specific responsibilities does not immediately affect the system's condition. As a result, the probabilities of human error should be modelled from the entire system or process perspectives. The total reliability should be assessed across a broader range by establishing a network of the individual human acts required to accomplish system goals. This is the method of determining human reliability. Additionally, in the modern system, procedures that rely on machines and software have surpassed duties entirely constituted by human acts. There is no use in analysing human reliability exclusively through the viewpoint of such a system. Thus, when the reliability of humans, machines, and software is correctly represented, as well as the reliability of mutual interference, it is feasible to evaluate the safety of a real system. Based on this fact, this study expands the approach from human error mode to human reliability assessment and system reliability assessment via extended human error evaluation in understanding human mechanisms that correlate to the level of complexity and interaction of the system.

The first chapter of the study addressed specific issues relating to human reliability assessment and its maritime applications, while the second chapter conducted critical literature reviews to identify research gaps. Following an in-depth evaluation of the literature, the next step is to establish research objectives and goals and the research structures necessary to accomplish them. Four case studies were conducted, each with a varying degree of complexity and circumstances. Prior to conducting case studies, Chapter 5 analysed numerous HRA methods to determine the most appropriate HRA strategy for each case study in this research. Chapter 6 developed a direct human error mode determination method based on BN-CREAM for estimating the overall probability of human error in a given context. The method was demonstrated during an emergency steering manoeuvre. Following that, a framework based on CREAM and Fuzzy theory was developed in Chapter 7 to anticipate all human failure probabilities during engine room fire-fighting procedures. Then, a novel approach based on a system reliability theory was introduced to examine the interaction between multiple human performance issues. The approach was demonstrated in Chapter 8 through an emergency preparedness case study involving a man-overboard situation. Finally, Chapter 9 provides the complete framework for system reliability assessment, which is proposed to assess system reliability, including human error, in a complex system of maritime operations. In addition, in Chapter 9, the framework for evaluating system reliability is applied to a case study, evaluating the system reliability in a ship-to-ship LNG bunkering operation.

10.3 Achievement of research aim and objectives

This PhD study aimed to develop a more consistent and comprehensive human reliability assessment framework that can extend to system reliability and can be applied to various maritime operations and systems. Second, the developed human reliability assessment framework should provide not only a single method but also different approaches optimised for various analytical ranges that differ according to the activity's complexity, data availability, human and material resources, and the significance of failure. Thus, the overall framework developed within this research study is focused on fulfilling the aforementioned aim. As a result of this study, four frameworks for assessing human reliability have been developed, each with its own set of optimised methods for research purposes, ranging from predicting human errors that may occur during simple procedures to analysing the system reliability of complex maritime systems.

Additionally, the following specific research objectives established in Chapter 3 were fulfilled as below:

- To critically review the literature relevant to the current maritime human error prediction, human reliability, and system safety to identify the shortcomings of the recent research and available methods.

An extensive critical review of human reliability assessment methods and research for practical applications (Chapter 2) was conducted. To begin, a review of the literature was undertaken to aid in the understanding of human error. Then, a further review of the overall process and framework for human reliability assessment was conducted. Following that, an in-depth evaluation of detailed HRA methods was performed, with specific attention to HRA-related research conducted maritime. Finally, section 2.8 derived and discussed the shortcomings and application issues associated with the previous HRA approach for identifying the research gap.

- To derive or develop customised performance shaping factors suitable for each context arising from the Maritime’s operation or system characteristics.

In Chapter 8, customised PSFs were derived for man-over-board scenarios, and guidance for rating each PSF was developed. Moreover, PSFs developed for context evaluation of the LNG bunkering situation introduced in Chapter 9 provide possible application forms and keywords, enabling specific evaluation criteria that existing HRA methods have not provided.

- To modify and develop advanced human error quantification techniques.

Four improved HRA approaches were presented in this study for quantifying human errors. To begin, the BN-CREAM technique was applied to Chapter 5's case study of emergency steering operation, which calculates the overall failure probability of human error by evaluating only the contextual factors related to a specific task. Following that, a case study on an engine room fire drill was conducted using the CREAM approach in conjunction with Fuzzy theory to quantify human errors in a more comprehensive approach dealing with specific person tasks, particularly those involving cognitive activity. Following that, a case study was performed for the man overboard drill using a customised PSF in conjunction with the SPAR-H technique. Finally, a new approach based on SLIM was proposed for quantifying human error in complicated structures such as LNG bunkering systems.

- Enhance human error identification techniques for a complex system.

To identify human errors, the human duties necessary to accomplish system goals were broken into the simplest elements possible through task analysis. The hierarchical task analysis technique was adopted to achieve this, and additional data were obtained using tabular task analysis. Additionally, STPA was proposed to analyse the interaction relationship between various system components within a complex system to discover human errors that could not be identified using existing approaches. STPA's primary objective is to identify and analyse human behaviours influencing the system risk associated with human-machine interaction. As a result, this study contributed to the development of a variety of human error detection algorithms appropriate for the complexity of a given research topic.

- Enhance human error representation modelling to assess system reliability and integrate errors into PRA

A human error representation model was developed to assist in the translation of human responsibilities into a system element for each activity. Reliability block diagrams were utilised for representing an entire system using individual human errors while taking their dependencies and system configurations into account to consider recovery actions and interactions among tasks. Chapters 8 and 9 demonstrate how the proposed method for modelling human error was used in case studies to assess human and system reliability.

- Establish criteria and a mechanism for determining the most appropriate HRA approach for a particular project.

An approach was developed to select an appropriate method for human reliability analysis from a pool of existing methods that satisfy the research objectives. Four criteria and associated fourteen sub-criteria for identifying the optimal method were defined as a guideline for evaluating HRA approaches. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method was utilised to prioritise the HRA techniques from a number of selected HRA methods. Four evaluations of the case studies that form this thesis were conducted.

- The following frameworks were developed to provide an optimum assessment approach for various purposes by considering the characteristics of the analysis target.

Finally, four frameworks based on enhanced HRA methodologies were constructed as primary research outputs, and their applications to case studies were tested to validate and show the frameworks' feasibility.

1) Develop an instant human error calculation model that responds to immediate and straightforward analysis needs. (Chapter 6)

2) Develop a human error calculation framework for extended human activities (Chapter 7)

3) Develop a human reliability assessment framework to integrate human error into a probabilistic risk assessment framework that can extend to system reliability assessment (Chapter 8)

4) Develop a human reliability assessment framework in a complex system by enhancing the human error identification approach (Chapter 9)

10.4 Limitations of the research study

Despite previous research efforts in the literature, it was identified that maritime researchers are still unable to understand the human interaction within the system adequately. There is an increasing number of variables to consider when it comes to the operation of the system at sea, which has become more complex in modern ship systems. In this setting, determining human reliability and the system's total reliability is a significant difficulty. As a result, numerous assumptions were made, which may have caused a loss of information and uncertainty. As a result, the limitations identified throughout this research study are as follows.

- As noted in Chapter 2, one of the primary challenges in HRA research is a lack of human error data. As a result, it was inevitable that the data shortage would inevitably arise over the course of this research. The normal human error data used in this research was generated from existing HRA procedures. However, there is an inherent limitation to accepting those nominal error probabilities that were not created with the maritime industry in mind but with the nuclear and aviation sectors. Even though the error value was adjusted using a PSF customised for the maritime sector to overcome these limitations, this study must admit the nominal human error value's underlying limitations.
- Along with data issues on human error, there is a lack of studies on human behaviour at sea. Even if a sequence of tasks comprising a system or operation is divided and extracted into a comprehensible minimal unit of human behaviour, the general human behaviour type corresponding to the understandable minimum unit of human behaviour is not formed in accordance with the maritime specifics. For instance, HEART's eight generic task types describe highly generalised and typical patterns of human behaviour, posing limitations in describing human behaviour throughout diverse maritime activities. Therefore, there is a limit to claiming that the types of human behaviour used in this study completely represent the types of human activities at sea.
- Despite numerous advancements and new investigations on PSFs in this study, it remains challenging to articulate the components that influence human performance.

This is also a natural constraint for developing PSF. PSF is difficult to quantify since separate elements do not act independently of one another in human behaviour. Additionally, the measurement of this PSF is complicated because it is inherently qualitative in nature, relying on expert judgement rather than quantitative measurement.

- The four case studies given in this paper were designed to illustrate a range of scenarios that may occur at sea. Nonetheless, the case examples offered do not encompass the whole spectrum of human behaviours that can occur at sea. Numerous complicated types of human behaviour can occur concurrently when performing tasks at sea, but the number and type of case studies examined were limited due to practical constraints.
- An expert-based approach was used to evaluate the operating system's condition at sea and generate necessary data. However, the number of experts who participated in this study was restricted, and experts' knowledge of human factors was limited.
- The situations or processes employed in the analysis were analysed progressively. However, the study does not represent a dynamic process that could occur in a repeating and reverse order.
- Data based on assumptions and literature were used to analyse the reliability of machines and software used to analyse system reliability. In complex systems, hardware and software reliability are also important factors, so research on this part could not be conducted because it went beyond the scope of this study. Still, this integration is required to evaluate overall system reliability.

10.5 Recommendations for future research

Based on the limitations given in the previous section, recommendations for future research are listed below:

- Human error data collecting is substantially more difficult in the maritime industry than in other industries. This is because the ship's operations are not limited to a certain place, and its operation pattern is complex. Therefore, research is required to establish a uniform platform that systematically collects human error data to combine data from various sources. For instance, efforts such as constructing online survey platforms and creating numerous Likert scales for easy-to-understand practical assessments are required to quantify and accurately reflect the expertise of a more significant number of experts. In addition, maritime operations will require a structured classification of human errors to prevent data duplication and omission. In addition, it is necessary to model the experimental design and execution of additional human error data because it is challenging to acquire objective data on specific human performance through observation.

- Human tasks onboard extend beyond ship manoeuvring. Consequently, the range of human activities at sea is extensive. Numerous procedures are required, including cargo handling, ballasting, anchoring, bunkering, and mooring. Moreover, each technique has its context and demands human behaviour. To adequately explain the details of maritime operations, identify distinct PSFs, and assess their influence, it is necessary to build a tailored nominal human error probability. Therefore, more particular PSFs must be found for the maritime industry, and their impact on human performance must be examined. To achieve this objective, establish more specific PSFs for the maritime industry. Second, more research is required to determine how PSF impacts human performance. Third, it is vital to develop quantitative or easily quantifiable PSFs to ensure consistency and objectivity in evaluations. To ensure consistency and objectivity in evaluations, measurable or easily quantifiable PSFs are required.
- Although current technologies offer substantial benefits to maritime operations, they also present new safety risks. For instance, the automation of ship operations, which is rapidly becoming a reality, no longer faces opposition. Thus, it is required to analyse the system reliability of new technology before implementing emerging technologies. Consequently, evolving technologies require fast action to address the increased hazards posed. Primarily, it is essential to investigate the various situational awareness generated by introducing a human-machine interface environment such as augmented reality and remote control stations.
- The study's scope should be broadened to include a variety of scenarios and procedures, ranging from ordinary onboard activities such as maintenance to emergency response abilities with critical operations. For example, it is essential to have systems that can adapt to rapidly advancing technologies, such as cybersecurity

and autonomous ships. This is because defining and categorising various types of human activity cannot be done independently; instead, it must be performed as part of the task analysis of certain maritime operations and systems. Additionally, case studies on more diverse maritime operations should be conducted to determine how implementation-related cognitive and physical acts may be related to actual sea activity.

- Finally, research is required to develop effective ways of implementing quantified human error in the maritime industry. Priority one should be research into incorporating human elements into the regulatory framework. The current formal safety assessment framework has no defined approach for addressing human factors. The subsequent step is to include an evaluation method for human factors in an individual safety system. Moreover, it should be studied how to facilitate end-user participation. For instance, if the results of risk analysis are to aid ship operations indeed, software aid like a dashboard that enables users to support choices, including the interpretation of research results actively, is required.

10.6 Chapter summary

This chapter discussed the research study and its uniqueness. Second, the research study's primary contributions were highlighted. Additionally, a summary of the aim and objectives of this research study was included, as well as the limitations. Finally, recommendations for future research were made.

11 Conclusion

11.1 Overview

This chapter presents a summary of the study's overall conclusions.

11.2 Concluding statements

This PhD study provides a flexible, ideal framework for evaluating the reliability of humans and systems, allowing them to be optimised in varied circumstances. However, there are different human errors, and the path from these errors to accidents is not intuitively apparent. Therefore, conceptual frameworks such as diagrams and models used to depict abstract processes for systematic recognition and classification of human error are essential for understanding the mechanism of human error. In this regard, explaining all phenomena using a single method is challenging. Consequently, human reliability assessments can be regarded similarly to identifying the specific degree of each object using a multi-measurement device that best matches the features of each object. Therefore, this study developed various human reliability assessment frameworks that may be applied to analysing diverse maritime systems and operations. As a result of these efforts, this research study contributed to developing an enhanced human reliability assessment framework for boosting human error identification, quantification, and modelling and suiting diverse analytical needs depending on the system's complexity and interaction. To demonstrate the framework's practical application and feasibility, they were applied to emergency response procedures for emergency steering, engine room fires, man overboard situations, and critical functions of the LNG bunkering process. Four(4) frameworks for human reliability assessment developed in this manner will contribute to the widespread usage of human reliability assessment, ranging from basic analyses performed by ship crew to in-depth analyses conducted by human reliability experts.

Therefore, this work contributed to five distinct research interests. First, the approach provided in this study for selecting the optimal human reliability analysis method broadens the scope of human reliability research in the maritime sector by combining the analysis objective with the most suitable analysis method. Choosing an ideal human reliability assessment method enables a thorough understanding of human factors, preventing the omission of the impact of human error on the system while generating cost-benefit research results. Next, the first case study applied the BN-CREAM to emergency steering operations. This study identified a control mode with an overall human error probability for emergency steering operations. In contrast to the present BN-CREAM model, which utilises a group of CPCs to reduce the effort required for calculation, the suggested method contains all the necessary individual logic to prevent data loss. In addition, the provision of a human error mode estimation model based on the BN-CREAM facilitates end-user access to human reliability analysis. Thirdly, the expanded CREAM methodology applied to the engine room fire drill demonstrates the likelihood of human failure for each human activity. The proposed strategy effectively takes into account the relative significance of each CPC. In addition, the use of evidence reasoning improves the calculation's precision. Fourthly, the SPAR-H framework provides a customised PSF list with accompanying guidelines. Unique reliability block diagrams depict estimated human errors to incorporate into overall system reliability. The presented method adequately indicates the effectiveness of redundancy and task dependency. Finally, using a novel human process model, the combination of STPA and SLIM identifies human responsibilities, with the system state revealing the path to the event scenario. A quantitative risk model can account for human error using the proposed method. Especially, a comparative analysis reveals an optimal design solution.

Although various risk analysis approaches have been applied to maritime scenarios, the results have not been integrated into probabilistic risk frameworks. Most evaluations are limited to qualitative analysis or include probability measurements of human error without proper modelling. Therefore, this study's methods and findings can be applied to the maritime industry to enhance the implementation of the HRA technique in the following manners. The offered methodologies and findings can be implemented differently concerning the probability of human error and the evaluation of PSF. First, the probability of human error can be utilised to identify and prioritise vulnerable areas in the ship's procedures and operational systems. Incorporating human factors into probabilistic risk assessment can also be used to estimate overall risk based on quantified human errors. The method can also be used to improve system design and procedures by analysing how the overall system's reliability changes due to new techniques and variations in how people collaborate. Moreover, when resources for risk analysis are restricted, like in the case of the captain's sole decision, an instant calculation model of human error can be implemented to estimate the ship's overall risk level to support the decision-making process of the ship. In the meantime, given that quantified human errors are the consequence of analysing the human task and the surrounding environment, the PSF evaluation results may help to identify and address vulnerable PSFs. For instance, the effect of human-machine interfaces on human performance can be assessed and included in a ship's user interface and layout design. This research will improve maritime safety by examining human errors at sea, identifying further problems, and adopting safety measures.

11.3 Chapter summary

This chapter summarised the author's conclusions on this research work.

Reference

ABBASPOUR, A., SAREMI, M., ALIBABAEI, A. & MOGHANLU, P. S. 2020. Determining the optimal human reliability analysis (HRA) method in healthcare systems using Fuzzy ANP and Fuzzy TOPSIS. *Journal of Patient Safety and Risk Management,* 25**,** 123-133.

ABBASSI, R., KHAN, F., GARANIYA, V., CHAI, S., CHIN, C. & HOSSAIN, K. A. 2015. An integrated method for human error probability assessment during the maintenance of offshore facilities. *Process Safety and Environmental Protection,* 94**,** 172-179.

ABS 2017. Guide for LNG Bunkering. ABS(American Bureau of Shipping), Houston, United States.

AHN, S. I. & KURT, R. E. 2020. Application of a CREAM based framework to assess human reliability in emergency response to engine room fires on ships. *Ocean Engineering,* 216**,** 108078.

AHN, S. I., KURT, R. E. & AKYUZ, E. 2022. Application of a SPAR-H based framework to assess human reliability during emergency response drill for man overboard on ships. *Ocean Engineering,* 251**,** 111089.

AINSWORTH, L. K. & KIRWAN, B. 1992. *A guide to task analysis*, Taylor & Francis.

AKYUZ, E. 2015. Quantification of human error probability towards the gas inerting process on-board crude oil tankers. *Safety Science,* 80**,** 77-86.

AKYUZ, E. 2016. Quantitative human error assessment during abandon ship procedures in maritime transportation. *Ocean Engineering,* 120**,** 21-29.

AKYUZ, E., CELIK, E. & CELIK, M. 2018. A practical application of human reliability assessment for operating procedures of the emergency fire pump at ship. *Ships and Offshore Structures,* 13**,** 208-216.

AKYUZ, E. & CELIK, M. 2014. Utilisation of cognitive map in modelling human error in marine accident analysis and prevention. *Safety science,* 70**,** 19-28.

AKYUZ, E. & CELIK, M. 2015. Application of CREAM human reliability model to cargo loading process of LPG tankers. *Journal of Loss Prevention in the Process Industries,* 34**,** 39-48.

AKYUZ, E. & CELIK, M. 2016. 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,* 43**,** 424-431.

AKYUZ, E., CELIK, M. & CEBI, S. 2016. A phase of comprehensive research to determine marine-specific EPC values in human error assessment and reduction technique. *Safety science,* 87**,** 63-75.

ANEZIRIS, O., KOROMILA, I. & NIVOLIANITOU, Z. 2020. A systematic literature review on LNG safety at ports. *Safety Science,* 124**,** 104595.

ANNETT, J. 2003. Hierarchical task analysis. *Handbook of cognitive task design,* 2**,** 17-35.

ANNETT, J. & STANTON, N. A. 2000. *Task analysis*, CRC Press.

ANTÃO, P. & SOARES, C. G. 2019. Analysis of the influence of human errors on the occurrence of coastal ship accidents in different wave conditions using Bayesian Belief Networks. *Accident Analysis & Prevention,* 133**,** 105262.

BAALISAMPANG, T., ABBASSI, R., GARANIYA, V., KHAN, F. & DADASHZADEH, M. 2018. Review and analysis of fire and explosion accidents in maritime transportation. *Ocean Engineering,* 158**,** 350-366.

BADOKHON, O. H. O. 2018. Development of a model for integrating resilience engineering principles to ship management system [sic] to enhance navigational bridge operation.

BALMAT, J.-F., LAFONT, F., MAIFRET, R. & PESSEL, N. 2011. A decision-making system to maritime risk assessment. *Ocean Engineering,* 38**,** 171-176.

BAYSARI, M. T., CAPONECCHIA, C., MCINTOSH, A. S. & WILSON, J. R. 2009. Classification of errors contributing to rail incidents and accidents: A comparison of two human error identification techniques. *Safety Science,* 47**,** 948-957.

BELL, J. & HOLROYD, J. 2009. Review of human reliability assessment methods. *Health & Safety Laboratory,* 78.

BELL, J. & WILLIAMS, J. Consolidation of the HEART human reliability assessment principles. Symposium Series No, 2016.

BLACKMAN, H. S., GERTMAN, D. I. & BORING, R. L. Human error quantification using performance shaping factors in the SPAR-H method. Proceedings of the human factors and ergonomics society annual meeting, 2008. SAGE Publications Sage CA: Los Angeles, CA, 1733-1737.

BOOKS, H. 2009. Reducing error and influencing behaviour. HSE.

BORING, R. L. 2010. How many performance shaping factors are necessary for human reliability analysis? : Idaho National Laboratory (INL).

BORING, R. L., GRIFFITH, C. D. & JOE, J. C. The measure of human error: Direct and indirect performance shaping factors. 2007 IEEE 8th Human Factors and Power Plants and HPRCT 13th Annual Meeting, 2007. IEEE, 170-176.

BYE, A., LAUMANN, K., TAYLOR, C., RASMUSSEN, M., ØIE, S., VAN DE MERWE, K., ØIEN, K., BORING, R., PALTRINIERI, N. & WÆRØ, I. 2017. The petro-HRA guideline.

CACCIABUE, P. C. 2004. Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training. *Reliability Engineering & System Safety,* 83**,** 229-240.

CALHOUN, J., SAVOIE, C., RANDOLPH-GIPS, M. & BOZKURT, I. 2014. Human reliability analysis in spaceflight applications, part 2: modified CREAM for spaceflight. *Quality and Reliability Engineering International,* 30**,** 3-12.

CASTIGLIA, F., GIARDINA, M. & TOMARCHIO, E. 2015. THERP and HEART integrated methodology for human error assessment. *Radiation Physics and Chemistry,* 116**,** 262-266.

CELIK, M. & CEBI, S. 2009. Analytical HFACS for investigating human errors in shipping accidents. *Accident Analysis & Prevention,* 41**,** 66-75.

CHAAL, M., BANDA, O. A. V., GLOMSRUD, J. A., BASNET, S., HIRDARIS, S. & KUJALA, P. 2020. A framework to model the STPA hierarchical control structure of an autonomous ship. *Safety Science,* 132**,** 104939.

CHAN, S. R., HAMID, N. A. & MOKHTAR, K. 2016. A theoretical review of human error in maritime accidents. *Advanced Science Letters,* 22**,** 2109-2112.

CHAUVIN, C. 2011. Human factors and maritime safety. *The Journal of Navigation,* 64**,** 625-632.

COOPER, S. E., RAMEY-SMITH, A., WREATHALL, J. & PARRY, G. 1996. A technique for human error analysis (ATHEANA). Nuclear Regulatory Commission.

DARBRA, R.-M. & CASAL, J. 2004. Historical analysis of accidents in seaports. *Safety science,* 42**,** 85-98.

DE FELICE, F. & PETRILLO, A. 2018. Human factors and reliability engineering for safety and security in critical infrastructures. *Decision Making, Theory, and Practice, Springer, Cham*.

DE MAYA, B. N., AHN, S. & KURT, R. Statistical analysis of MAIB database for the period 1990-2016. 18th International Congress of the Maritime Association of the Mediterranean, 2019a.

DE MAYA, B. N., BABALEYE, A. O. & KURT, R. E. 2019b. Marine accident learning with fuzzy cognitive maps (MALFCMs) and Bayesian networks. *Safety in Extreme Environments*, 1-10.

DEKKER, S. 2017a. *The field guide to human error investigations*, Routledge.

DEKKER, S. 2017b. *The field guide to understanding 'human error'*, CRC press.

DGHAYM, D., HOANG, T. S., TURNOCK, S. R., BUTLER, M., DOWNES, J. & PRITCHARD, B. 2021. An STPA-based formal composition framework for trustworthy autonomous maritime systems. *Safety science,* 136**,** 105139.

DIAPER, D. & STANTON, N. 2003. The handbook of task analysis for human-computer interaction.

ELIOPOULOU, E., PAPANIKOLAOU, A. & VOULGARELLIS, M. 2016. Statistical analysis of ship accidents and review of safety level. *Safety science,* 85**,** 282-292.

ELMARAGHY, W. H., NADA, O. & ELMARAGHY, H. A. 2008. Quality prediction for reconfigurable manufacturing systems via human error modelling. *International Journal of Computer Integrated Manufacturing,* 21**,** 584-598.

EMBREY, D. 2004. Qualitative and quantitative evaluation of human error in risk assessment. *Human factors for engineers,* 2**,** 151.

EMBREY, D., HUMPHREYS, P., ROSA, E., KIRWAN, B. & REA, K. 1984. SLIM-MAUD: an approach to assessing human error probabilities using structured expert judgment. Volume I. Overview of SLIM-MAUD. Brookhaven National Lab., Upton, NY (USA).

EMSA 2018. Guidance on LNG Bunkering to Port Authorities and Administration. EMSA (European Maritime Safety Agency) Lisbon, Portugal.

FAHSSI, R., MARTINIE, C. & PALANQUE, P. Enhanced task modelling for systematic identification and explicit representation of human errors. IFIP Conference on Human-Computer Interaction, 2015. Springer, 192-212.

FERDOUS, R., KHAN, F., SADIQ, R., AMYOTTE, P. & VEITCH, B. 2009. Handling data uncertainties in event tree analysis. *Process safety and environmental protection,* 87**,** 283-292.

FUJITA, Y. & HOLLNAGEL, E. 2004. Failures without errors: quantification of context in HRA. *Reliability Engineering & System Safety,* 83**,** 145-151.

GERTMAN, D., BLACKMAN, H., MARBLE, J., BYERS, J. & SMITH, C. 2005. The SPAR-H human reliability analysis method. *US Nuclear Regulatory Commission,* 230**,** 35.

GIL, M., WRÓBEL, K. & MONTEWKA, J. 2019. Toward a method evaluating control actions in STPA-based model of ship-ship collision avoidance process. *Journal of Offshore Mechanics and Arctic Engineering,* 141.

GOSLING, G., ROBERTS, K. H. & JAYASWAL, A. 1998. Improving the Representation of Human Error in the Use of the Flight Crew Human Factors Integration Tool.

HÄNNINEN, M. 2014. Bayesian networks for maritime traffic accident prevention: benefits and challenges. *Accident Analysis & Prevention,* 73**,** 305-312.

HE, X., WANG, Y., SHEN, Z. & HUANG, X. 2008. A simplified CREAM prospective quantification process and its application. *Reliability Engineering & System Safety,* 93**,** 298-306.

HOLLNAGEL, E. 1998. *Cognitive reliability and error analysis method (CREAM)*, Elsevier.

HOLLNAGEL, E. 2005. Human reliability assessment in context. *Nuclear engineering and technology,* 37**,** 159-166.

HOLLNAGEL, E. 2017a. *FRAM: the functional resonance analysis method: modelling complex socio-technical systems*, CRC Press.

HOLLNAGEL, E. 2017b. *Safety-II in practice: developing the resilience potentials*, Taylor & Francis.

IACS 2017. LNG Bunkering Guidelines, Second Edition, Rec142. IACS(International Association of Classification Societies).

IMO 2001. the Regulation 19 of SOLAS Chapter III.

IMO 2018a. MARPOL ANNEX IV. *In:* ORGANIZATION(IMO), I. M. (ed.).

IMO, S. C. E. 2018b. IMO. London.

ISLAM, R., ABBASSI, R., GARANIYA, V. & KHAN, F. 2017a. Development of a human reliability assessment technique for the maintenance procedures of marine and offshore operations. *Journal of Loss Prevention in the Process Industries,* 50**,** 416-428.

ISLAM, R., ABBASSI, R., GARANIYA, V. & KHAN, F. I. 2016. Determination of human error probabilities for the maintenance operations of marine engines. *Journal of Ship Production and Design,* 32**,** 226-234.

ISLAM, R., KHAN, F., ABBASSI, R. & GARANIYA, V. 2018. Human error probability assessment during maintenance activities of marine systems. *Safety and health at work,* 9**,** 42-52.

ISLAM, R., YU, H., ABBASSI, R., GARANIYA, V. & KHAN, F. 2017b. Development of a monograph for human error likelihood assessment in marine operations. *Safety science,* 91**,** 33-39.

JAHANGIRI, M., HOBOUBI, N., ROSTAMABADI, A., KESHAVARZI, S. & HOSSEINI, A. A. 2016. Human error analysis in a permit to work system: a case study in a chemical plant. *Safety and health at work,* 7**,** 6-11.

JANG, H., JEONG, B., ZHOU, P., HA, S. & NAM, D. 2021. Demystifying the lifecycle environmental benefits and harms of LNG as marine fuel. *Applied Energy,* 292**,** 116869.

JIAN-BO, Y. & DONG-LING, X. 2002. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* 32**,** 289-304.

KAMYAB, S., NEMATOLLAHI, M. & SHAFIEE, G. 2013. Sensitivity analysis on the effect of software-induced common cause failure probability in the computer-based reactor trip system unavailability. *Annals of Nuclear Energy,* 57**,** 294-303.

KANG, H. G., LIM, H. G., LEE, H. J., KIM, M. C. & JANG, S. C. 2009. Input-profile-based software failure probability quantification for safety signal generation systems. *Reliability Engineering & System Safety,* 94**,** 1542-1546.

KAYISOGLU, G., GUNES, B. & BESIKCI, E. B. 2021. SLIM based Methodology for Human Error Probability Calculation of Bunker Spills In Maritime Operations. *Reliability Engineering & System Safety***,** 108052.

KHALAQUZZAMAN, M., LEE, S. J., KIM, M. C. & JUNG, W. 2014. Estimation of reactor protection system software failure probability considering undetected faults. *Nuclear Engineering and Design,* 280**,** 201-209.

KHAN, F. 2008. Human factors special issue. *Journal of Loss Prevention in the Process Industries,* 21**,** 225-226.

KIRWAN, B. 1992. Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied ergonomics,* 23**,** 299-318.

KIRWAN, B. 1994. *A guide to practical human reliability assessment*, CRC press.

KIRWAN, B. 1996. The validation of three human reliability quantification techniques—THERP, HEART and JHEDI: Part 1—technique descriptions and validation issues. *Applied ergonomics,* 27**,** 359-373.

KIRWAN, B. 1998. Human error identification techniques for risk assessment of high risk systems—Part 1: review and evaluation of techniques. *Applied ergonomics,* 29**,** 157-177.

KIRWAN, B. & AINSWORTH, L. K. 1992. *A guide to task analysis: the task analysis working group*, CRC press.

KIRWAN, B., MARTIN, B., RYCRAFT, H. & SMITH, A. 1990. Human error data collection and data generation. *International Journal of Quality & Reliability Management*.

KLAW. 2022. Available: https://www.klawlng.com/ [Accessed 0222 2022].

KONSTANDINIDOU, M., NIVOLIANITOU, Z., KIRANOUDIS, C. & MARKATOS, N. 2006. A fuzzy modeling application of CREAM methodology for human reliability analysis. *Reliability Engineering & System Safety,* 91**,** 706-716.

KRISTIANSEN, S. 2013. *Maritime transportation: safety management and risk analysis*, Routledge.

KURT, R., ARSLAN, V., KHALID, H., COMRIE, E., BOULOUGOURIS, E. & TURAN, O. SEAHORSE procedure improvement system: development of instrument. International SEAHORSE Conference on Maritime Safety and Human Factors, 2016a. 1-8.

KURT, R., ARSLAN, V., TURAN, O., DE WOLFF, L., WOOD, B., ARSLAN, O., KECECI, T., WINKELMAN, J., VAN WIJNGAARDEN, M. & PAPADAKIS, G. 2015. SEAHORSE project: Dealing with maritime workarounds and developing smarter procedures.

KURT, R. E., KHALID, H., TURAN, O., HOUBEN, M., BOS, J. & HELVACIOGLU, I. H. 2016b. Towards human-oriented norms: Considering the effects of noise exposure on board ships. *Ocean Engineering,* 120**,** 101-107.

LACEY, P. An application of fault tree analysis to the identification and management of risks in government funded human service delivery. Proceedings of the 2nd International Conference on Public Policy and Social Sciences held in Kuching, 2011.

LEE, J. & CHUNG, H. 2018. A new methodology for accident analysis with human and system interaction based on FRAM: Case studies in maritime domain. *Safety science,* 109**,** 57-66.

LEE, S. M., HA, J. S. & SEONG, P. H. 2011a. CREAM-based communication error analysis method (CEAM) for nuclear power plant operators' communication. *Journal of Loss Prevention in the Process Industries,* 24**,** 90-97.

LEE, S. W., KIM, A. R., HA, J. S. & SEONG, P. H. 2011b. Development of a qualitative evaluation framework for performance shaping factors (PSFs) in advanced MCR HRA. *Annals of Nuclear Energy,* 38**,** 1751-1759.

LEVESON, N. 2004. A new accident model for engineering safer systems. *Safety science,* 42**,** 237-270.

LEVESON, N. G. & THOMAS, J. P. 2018. STPA handbook. *NANCY LEVESON AND JOHN THOMAS,* 3**,** 1-188.

LI, J. & HUANG, Z. 2012. Fire and explosion risk analysis and evaluation for LNG ships. *Procedia engineering,* 45**,** 70-76.

LIU, J., AYDIN, M., AKYUZ, E., ARSLAN, O., UFLAZ, E., KURT, R. E. & TURAN, O. 2021. Prediction of human–machine interface (HMI) operational errors for maritime autonomous surface ships (MASS). *Journal of Marine Science and Technology***,** 1-14.

LIU, P. & LI, Z. 2014. Human error data collection and comparison with predictions by SPAR-H. *Risk Analysis,* 34**,** 1706-1719.

MARINE-INSIGHT. 2021. Available: https://www.marineinsight.com/naval-architecture/understanding-steering-gear-ships/ [Accessed].

MARSEGUERRA, M., ZIO, E. & LIBRIZZI, M. 2007. Human Reliability Analysis by Fuzzy "CREAM". *Risk Analysis,* 27**,** 137-154.

MARTINS, M. R. & MATURANA, M. C. 2010. Human error contribution in collision and grounding of oil tankers. *Risk Analysis: An International Journal,* 30**,** 674-698.

MOSLEH, A., CHANG, J., CHANDLER, F., MARBLE, J., BORING, R. & GERTMAN, D. 2006. Human reliability analysis methods-selection guidance for nasa. *NASA, Tech. Rep*.

MUSHARRAF, M., HASSAN, J., KHAN, F., VEITCH, B., MACKINNON, S. & IMTIAZ, S. 2013. Human reliability assessment during offshore emergency conditions. *Safety science,* 59**,** 19-27.

NORMAN, D. A. 1983. Design rules based on analyses of human error. *Communications of the ACM,* 26**,** 254-258.

NOROOZI, A., KHAN, F., MACKINNON, S., AMYOTTE, P. & DEACON, T. 2014. Determination of human error probabilities in maintenance procedures of a pump. *Process Safety and Environmental Protection,* 92**,** 131-141.

ÖLÇER, A. & ODABAŞI, A. 2005. A new fuzzy multiple attributive group decision making methodology and its application to propulsion/manoeuvring system selection problem. *European Journal of Operational Research,* 166**,** 93-114.

PARHIZKAR, T., UTNE, I. B., VINNEM, J. E. & MOSLEH, A. 2021. Dynamic probabilistic risk assessment of decision-making in emergencies for complex systems, case study: Dynamic positioning drilling unit. *Ocean Engineering,* 237**,** 109653.

PARK, K. S. & IN LEE, J. 2008. A new method for estimating human error probabilities: AHP–SLIM. *Reliability Engineering & System Safety,* 93**,** 578-587.

PARK, M. 2018. Quality improvement in pediatric care. *Korean Journal of Pediatrics,* 61**,** 1-5.

PRAETORIUS, G. 2014. *Vessel Traffic Service (VTS): a maritime information service or traffic control system?: understanding everyday performance and resilience in a socio-technical system under change.* Chalmers tekniska högskola.

PRAETORIUS, G., GRAZIANO, A., SCHRÖDER-HINRICHS, J.-U. & BALDAUF, M. 2017. Fram in FSA—Introducing a function-based approach to the formal safety assessment framework. *Advances in Human Aspects of Transportation.* Springer.

QIAO, W., MA, X., LIU, Y. & DENG, W. 2022. Resilience evaluation of maritime liquid cargo emergency response by integrating FRAM and a BN: A case study of a propylene leakage emergency scenario. *Ocean Engineering,* 247**,** 110584.

REASON, J. 1990. *Human error*, Cambridge university press.

RIGBY, C. C. & FRANKS, L. 1970. A human tissue culture cell line from a transitional cell tumour of the urinary bladder: growth, chromosome pattern and ultrastructure. *British journal of cancer,* 24**,** 746-754.

ROKSETH, B., UTNE, I. B. & VINNEM, J. E. 2017. A systems approach to risk analysis of maritime operations. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability,* 231**,** 53-68.

ROSS, T. W. 2006. Ship’s lifeboats: Analysis of accident cause and effect and its relationship to seafarers’ hazard perception. Master's, Dalhousie University). Retrieved from http://www. msc1206. com ….

ROTHBLUM, A. M. Human error and marine safety. National Safety Council Congress and Expo, Orlando, FL, 2000.

SALAS, E., DRISKELL, J. E. & HUGHES, S. 1996. The study of stress and human performance. *Stress and human performance(A 97-27090 06-53), Mahwah, NJ, Lawrence Erlbaum Associates, Publishers, 1996*, 1-45.

SALIHOGLU, E. & BEŞIKÇI, E. B. 2021. The use of Functional Resonance Analysis Method (FRAM) in a maritime accident: A case study of Prestige. *Ocean Engineering,* 219, 108223.

SALMON, P. M., CORNELISSEN, M. & TROTTER, M. J. 2012. Systems-based accident analysis methods: a comparison of Accimap, HFACS, and STAMP. *Safety science,* 50, 1158-1170.

SCHRÖDER-HINRICHS, J.-U., HOLLNAGEL, E., BALDAUF, M., HOFMANN, S. & KATARIA, A. 2013. Maritime human factors and IMO policy. *Maritime Policy & Management,* 40, 243-260.

SCHRÖDER-HINRICHS, J. U., BALDAUF, M. & GHIRXI, K. T. 2011. Accident investigation reporting deficiencies related to organizational factors in machinery space fires and explosions. *Accident Analysis & Prevention,* 43, 1187-1196.

SENDERS, J. W. & MORAY, N. P. 1991. Human Error: Cause. *Prediction, and Reduction*.

SENDERS, J. W. & MORAY, N. P. 2020. *Human Error:: Cause, Prediction, and Reduction*, CRC Press.

SGOBBA, T. 2017. *Space Safety and Human Performance*, Butterworth-Heinemann.

SHAPPELL, S. A. & WIEGMANN, D. A. 2000. The human factors analysis and classification system--HFACS.

SHEPHERD, A. 1985. Hierarchical task analysis and training decisions. *PLET: Programmed Learning & Educational Technology,* 22, 162-176.

SHEPHERD, A. 2003. *Hierarchial task analysis*, Crc Press.

SHORROCK, S. T. & KIRWAN, B. 2002. Development and application of a human error identification tool for air traffic control. *Applied ergonomics,* 33, 319-336.

STANTON, N. & BABER, C. 1996. A systems approach to human error identification. *Safety science,* 22, 215-228.

STANTON, N. A. 2006. Hierarchical task analysis: Developments, applications, and extensions. *Applied ergonomics,* 37, 55-79.

SULTANA, S., OKOH, P., HAUGEN, S. & VINNEM, J. E. 2019. Hazard analysis: Application of STPA to ship-to-ship transfer of LNG. *Journal of Loss Prevention in the Process Industries,* 60**,** 241-252.

SUN, B., GUO, K. & PAREEK, V. K. 2014. Computational fluid dynamics simulation of LNG pool fire radiation for hazard analysis. *Journal of Loss Prevention in the Process Industries,* 29**,** 92-102.

SVENSON, O. 1989. On expert judgements in safety analyses in the process industries. *Reliability Engineering & System Safety,* 25**,** 219-256.

SWAIN, A. D. 1964. THERP. Sandia Corp., Albuquerque, N. Mex.

SWAIN, A. D. 1983. Handbook of human reliability analysis with emphasis on nuclear power plant applications. *NUREG/CR-1278, SAND 80-0200*.

SWAIN, A. D. & GUTTMANN, H. E. 1983. Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report. Sandia National Labs.

TAKAGI, T. & SUGENO, M. 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics***,** 116-132.

TANG, J. X., BAO, Y. K., WANG, L. C., GUO, C. X., LIU, W. H. & WANG, T. P. An application of CREAM for human reliability analysis in power system switching operation. Applied Mechanics and Materials, 2014. Trans Tech Publ, 2585-2588.

UFLAZ, E., CELIK, E., AYDIN, M., ERDEM, P., AKYUZ, E., ARSLAN, O., KURT, R. E. & TURAN, O. 2022. An extended human reliability analysing under fuzzy logic environment for ship navigation. *Australian Journal of Maritime & Ocean Affairs***,** 1-21.

UNG, S.-T. 2015. A weighted CREAM model for maritime human reliability analysis. *Safety science,* 72**,** 144-152.

UNG, S.-T. 2019. Evaluation of human error contribution to oil tanker collision using fault tree analysis and modified fuzzy Bayesian Network based CREAM. *Ocean Engineering,* 179**,** 159-172.

UNG, S. T. & SHEN, W. M. 2011. A novel human error probability assessment using fuzzy modeling. *Risk Analysis: An International Journal,* 31**,** 745-757.

VAGIAS, N. 2010. A bayesian network application for the prediction of human fatigue in the marine industry. *Unpublished dissertation, National Technical University of Athens*.

VOGT, W. P. & JOHNSON, B. 2011. *Dictionary of statistics & methodology: A nontechnical guide for the social sciences*, Sage.

WENG, J. & YANG, D. 2015. Investigation of shipping accident injury severity and mortality. *Accident Analysis & Prevention,* 76**,** 92-101.

WHALEY, A. M., KELLY, D. L., BORING, R. L. & GALYEAN, W. J. 2011. SPAR-H step-by-step guidance. *Idaho Falls (ID): Risk, Reliability, and NRC Programs Department, Idaho National Laboratory*.

WILLIAMS, J. HEART—a proposed method for achieving high reliability in process operation by means of human factors engineering technology. Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society 1985. Taylor & Francis, 5/1-5/15.

WOLT, J. D., KEESE, P., RAYBOULD, A., FITZPATRICK, J. W., BURACHIK, M., GRAY, A., OLIN, S. S., SCHIEMANN, J., SEARS, M. & WU, F. 2010. Problem formulation in the environmental risk assessment for genetically modified plants. *Transgenic research,* 19**,** 425-436.

WOODS, D. D., DEKKER, S., COOK, R., JOHANNESEN, L. & SARTER, N. 2017. *Behind human error*, CRC Press.

WU, B., YAN, X., WANG, Y. & SOARES, C. G. 2017. An evidential reasoning-based CREAM to human reliability analysis in maritime accident process. *Risk analysis,* 37**,** 1936-1957.

WU, B., YIP, T. L., YAN, X. & SOARES, C. G. 2019. Fuzzy logic based approach for ship-bridge collision alert system. *Ocean Engineering,* 187**,** 106152.

XI, Y., YANG, Z., FANG, Q., CHEN, W. & WANG, J. 2017. A new hybrid approach to human error probability quantification–applications in maritime operations. *Ocean Engineering,* 138**,** 45-54.

YANG, Z., ABUJAAFAR, K. M., QU, Z., WANG, J., NAZIR, S. & WAN, C. 2019. Use of evidential reasoning for eliciting bayesian subjective probabilities in human reliability analysis: A maritime case. *Ocean Engineering,* 186**,** 106095.

YANG, Z., BONSALL, S., WALL, A., WANG, J. & USMAN, M. 2013. A modified CREAM to human reliability quantification in marine engineering. *Ocean engineering,* 58**,** 293-303.

YOON, K. P. & HWANG, C.-L. 1995. *Multiple attribute decision making: an introduction*, Sage publications.

ZHANG, D., ZHANG, M., YAO, H., KAI, Z. & FAN, C. Prediction Model of Human Error Probability in Autonomous Cargo Ships. International Seminar on Safety and

Security of Autonomous Vessels (ISSAV) and European STAMP Workshop and Conference (ESWC), 2020a.

ZHANG, M., ZHANG, D., GOERLANDT, F., YAN, X. & KUJALA, P. 2019. Use of HFACS and fault tree model for collision risk factors analysis of icebreaker assistance in ice-covered waters. *Safety science,* 111**,** 128-143.

ZHANG, M., ZHANG, D., YAO, H. & ZHANG, K. 2020b. A probabilistic model of human error assessment for autonomous cargo ships focusing on human–autonomy collaboration. *Safety science,* 130**,** 104838.

ZHIQIANG, S., HONGWE, X., XIUJIAN, S. & FENGQIANG, L. 2009. Engineering approach for human error probability quantification. *Journal of Systems Engineering and Electronics,* 20**,** 1144-1152.

ZHOU, Q., WONG, Y. D., XU, H., VAN THAI, V., LOH, H. S. & YUEN, K. F. 2017a. An enhanced CREAM with stakeholder-graded protocols for tanker shipping safety application. *Safety science,* 95**,** 140-147.

ZHOU, T., WU, C., ZHANG, J. & ZHANG, D. 2017b. Incorporating CREAM and MCS into fault tree analysis of LNG carrier spill accidents. *Safety science,* 96**,** 183-191.

Appendix A

Table AA. 1 Tabular task analysis for rescue boat drill

Sequential tasks	Responsible person	Location	Interface or equipment	Required procedures or manuals
1. Detect a man overboard, report and move to a duty position				
1.1 Detect a man overboard	Watchkeeping crew	Wing bridge	N/A	N/A
1.2 Throw a life buoy into the sea	Watchkeeping crew	Wing bridge	Quick release lifebuoy	Manual for quick release lifebuoy
1.3 Report to a duty officer	Watchkeeping crew	Navigation bridge	N/A	N/A
1.4 Slow down speed and stop the ship	Duty officer	Navigation bridge	N/A	N/A
1.5 Mark the position where a man overboard	Duty officer	Navigation bridge	Telephone	N/A
1.6 Report to Master	Duty officer	Navigation bridge	Engine telegraph	N/A
1.7 Push the alarm and make an announcement	Duty officer	Navigation bridge	Alarm system, public address system	Manual for alarm system, Manual for PA system
1.8 Report to Authorities	Duty officer	Navigation bridge	Radio equipment	Manual for radio equipment and contact details
1.9 Take personal equipment	All rescue crews	Cabin room	Personal equipment	N/A
1.10 Move to muster station	All rescue crews	Cabin room	N/A	N/A
1.11 Move to Navigation bridge	All rescue crews	Accommodation	N/A	N/A
2. Ship manoeuvring to return to the position				
2.1 Manoeuvring a ship to where a man overboard	Master	Navigation bridge	Ship navigation equipment	Williamson or Anderson/ single turning manual

2.2 Search a man overboard	Master	Navigation bridge	N/A	N/A
2.3 Stop the ship	Master	Navigation bridge	N/A	N/A
3. Gathering and safety briefing at the muster station				
3.1 Wear personal equipment	All rescue crews	Muster station	Personal equipment	Lifejacket or immersion suit wearing instruction
3.2 Check the number of crews and their equipment	Chief officer	Muster station	Personal equipment	Lifejacket or immersion suit wearing instruction
3.3 Deliver a safety briefing	Chief officer	Muster station	N/A	N/A
4. Preparation of boat lowering and embarkation				
4.1 Separate the charging cable	Crew-A	Boat deck	Charging cable and connecter	Manual for davit operation
4.2 Close boat drain valves	Crew-A	Boat deck	Boat lashing wire	Manual for davit operation
4.3 Secure a painter line to the rescue boat painter hook	Crew-B	Boat deck	lashing stopper	Manual for davit operation
4.4 Remove boat securing wires	Crew-C	Boat deck	boat bottom plug	Manual for davit operation
4.5 Remove the lashing stopper on boat davit	Crew-C	Boat deck	Painter line and hook	Manual for davit operation
4.6 Rescue boat boarding	Crew-A, B, C	Boat deck	N/A	N/A
5. Boat testing and lowering				
5.1 Push the engine start button	Crew-A	Boat in a stowed position	Rescue boat engine	Manual for rescue boat operation
5.2 Move the steering wheel	Crew-A	Boat in a stowed position	Rescue boat steering gear	Manual for rescue boat operation

5.3 Pull out the safety pin for winch braker	Crew-D	Boat deck	safety pin for winch braker	Manual for davit operation
5.4 Raise the lever of the winch brake to lower boat to the sea level	Crew-D	Boat deck	winch brake, rescue boat davit	Manual for davit operation
6. Boat release				
6.1Check hydrostatic interlock lever position	Crew-A	Boat hanging condition	Hydrostatic interlock system	Manual for a boat release mechanism
6.2 On load release rescue boat				
6.2.1 pull out the release handle safety pin	Crew-A	Boat on the sea with fall	release handle safety pin	Manual for a boat release mechanism
6.2.2 Pull the release hand	Crew-A	Boat on the sea with fall	Boat release system	Manual for a boat release mechanism
6.2.3 Pull the painter release handle	Crew-A	Boat on the sea	Rescue boat painter release device	
6.3 Off load release rescue boat				
6.3.1 Pull out the release handle safety pin	Crew-A	Boat on the sea with fall	release handle safety pin	Manual for a boat release mechanism
6.3.2 Remove the hydrostatic interlock cover	Crew-A	Boat on the sea with fall	the hydrostatic interlock cover	Manual for a boat release mechanism
6.3.3 Lift up the hydrostatic interlock lever	Crew-A	Boat on the sea with fall	the hydrostatic interlock lever	Manual for a boat release mechanism
6.3.4 Pull the release hand	Crew-A	Boat on the sea with fall	Boat release system	Manual for a boat release mechanism
6.3.5 Open the boat's F & A hatch doors	Crew-B, C	Boat on the sea	Boat hatch door	Manual for a boat release mechanism
6.3.6 Disconnect suspension links from hook	Crew-B, C	Boat on the sea	Suspension links	Manual for a boat

				release mechanism
6.3.7 Pull the painter release handle	Crew-B	Boat on the sea	Painter release hook	Manual for rescue boat operation
7. Manoeuvring and rescue activity				
7.1 Manoeuvring the rescue boat to approach a man	Crew-A	Boat on the sea	Boat control system	Manual for rescue boat operation
7.2 Rescue a man (Pull up a man to a boat)	Crew-B, C	Boat on the sea		
7.3 Manoeuvring the rescue boat to the ship for recovery	Crew-A	Boat on the sea	Boat control system	Manual for rescue boat operation
8. Boat recovery				
8.1 Connect a painter line to the painter hook	Crew-B	Boat on the sea	painter and painter hook	Manual for rescue boat operation
8.2 Place FWD & AFT hooks in reset position simultaneously	Crew-B, C	Boat on the sea	FWD & AFT hooks and boat release system	Manual for a boat recovery mechanism
8.3 Push the release handle	Crew-A	Boat on the sea	Boat release system	Manual for a boat recovery mechanism
8.4 Insert the release handle safety pin	Crew-A	Boat on the sea	release handle safety pin	Manual for a boat recovery mechanism
8.5 Connect suspension links to boat release hooks	Crew-B, C	Boat on the sea	Suspension links and boat release hook	Manual for a boat recovery mechanism
8.6 Push the winch up button to stowed position	Crew-D	Boat on the sea	Boat release system	Manual for davit operation

Appendix B

List of Unsafety Control Actions for ESD system during LNG STS bunkering

Table AB. 1 Control Action 1. Setting up user configurable parameters in the system

Not providing Cause hazard	Providing Causes Hazard	Too early/ late
UCA1.1 Operator did not set up system user configurable parameter for max gas detection level in cargo machinery space. As a result, when gas detection in cargo machinery space was above 60% LEL, ESD was not activated automatically by the ESD system. (H1)	UCA1.2 Operator did wrongly set up system user configurable parameter for max gas detection level in cargo machinery space. As a result, when gas detection in cargo machinery space was above 60% LEL, ESD was not activated automatically by the ESD system. (H1)	N/A
UCA1.3 Operator did not set up system user configurable parameter for max gas detection level at bunkering manifold area. As a result, when gas detection at bunkering manifold area above 60% LEL, ESD was not activated automatically by the ESD system. (H2)	UCA1.4 Operator did wrongly set up system user configurable parameter for max gas detection level at bunkering manifold area. As a result, when gas detection at bunkering manifold area above 60% LEL, ESD was not activated automatically by the ESD system.	N/A
UCA1.5 Operator did not set up system user configurable parameter for max pressure at bunkering manifold. As a result, when the pressure at the bunkering manifold is high, ESD was not automatically activated by the ESD system.	UCA1.6 Operator wrongly set up system user configurable parameter for max pressure at bunkering manifold. As a result, when the pressure at the bunkering manifold is high, ESD was not automatically activated by the ESD system.	N/A
UCA1.7 Operator did not set up system user configurable parameter for max pressure in the vapour return line. As a result, when the pressure in the vapour return line was high, ESD was not activated automatically by the ESD system.	UCA1.8 Operator wrongly set up system user configurable parameter for max pressure in the vapour return line. As a result, when the pressure in the vapour return line was high, ESD was not activated automatically by the ESD system.	N/A
UCA1.9 Operator did not set up system user configurable parameter for the event of ESD valve motive power loss. As a result, when ESD valve motive power was lost, ESD was not activated automatically by the ESD system.	N/A	N/A

UCA1.10 Operator did not set up system user configurable parameter for the event of ERC activation by fault. As a result, When ERC was activated by default, ESD was not activated automatically by the ESD system.	N/A	N/A
UCA1.11 Operator did not set up system user configurable parameter for the event of High-high liquid level in receiving tank. As a result, ESD was not automatically activated by the ESD system when the high-high liquid level was in the LNG receiving tank.	N/A	N/A
UCA1.12 Operator did not set up system user configurable parameter for the event of high-high pressure in LNG receiving tank. As a result, ESD was not automatically activated by the ESD system when high-high pressure was in the LNG receiving tank.	N/A	N/A
UCA1.13 Operator did not set up system user configurable parameter for the event of ESD signal receiving from receiving vessel. As a result, when ESD signal was received from receiving vessel, ESD was not activated automatically by the ESD system.	N/A	N/A
UCA1.14 Operator did not set up system user configurable parameter for the event of fire onboard. As a result, when fire onboard was detected, ESD was not activated automatically by the ESD system.	N/A	N/A

Table AB. 2 Control Action 2. Activate ESD manually

Not providing Cause hazard	Providing Causes Hazard	Timing error Too early/ late
UCA2.1 Operator did not activate ESD manually when gas detection in cargo machinery space above 60% LEL; other controllers activated the ESD. (H1)	UCA2.2 Operator did activate the ESD manually when gas detection in cargo machinery was normal. (H1)	UCA2.3 Operator did activate the ESD manually too late when other controllers did not activate gas detection in cargo machinery space above 60% LEL and ESD. (H1)
UCA2.4 Operator did not activate the ESD manually when other controllers did not activate gas detection at bunkering manifold area above 60% LEL and ESD. (H2)	UCA2.5 Operator did activate ESD manually when gas detection at the bunkering manifold area was normal. (H2)	UCA2.6 Operator did activate the ESD manually too late when other controllers did not activate gas detection at bunkering manifold area above 60% LEL and ESD. (H2)
UCA2.7 Operator did not activate the ESD manually when the pressure at the bunkering manifold was high, and other controllers did not activate the ESD. (H3)	UCA2.8 Operator did activate ESD manually when the pressure at the bunkering manifold was normal. (H3)	UCA2.9 Operator did activate ESD manually too late when the pressure at the bunkering manifold was high, and other controllers did not activate the ESD. (H3)
UCA2.10 Operator did not activate the ESD manually when the pressure in the vapour return line was high, and the ESD was not activated automatically by the ESD system.	UCA2.11 Operator did activate ESD manually when the pressure in the vapour return line was normal.	UCA2.12 Operator did activate the ESD manually too late when the pressure in the vapour return line was high, and ESD was not activated automatically by the ESD system.
UCA2.13 Operator did not activate the ESD manually when the ESD valve motive power was lost, and the ESD was not activated automatically by the ESD system.	N/A	UCA2.14 Operator did activate the ESD manually too late, the ESD valve motive power wass lost, and theESD was not activated automatically by the ESD system.
UCA2.15 Operator did not activate the ESD manually when the ERC was activated by default and the ESD was not activated automatically by the ESD system.	N/A	UCA2.16 Operator did activate the ESD manually too late, the ERC was activated by default, and the ESD was not activated automatically by the ESD system.

UCA2.17 Operator did not activate the ESD manually when the ESD system had a high-high liquid level in the LNG receiving tank, and the ESD was not activated automatically.	N/A	UCA2.18 Operator did activate the ESD manually too late when a High-high liquid level in the LNG receiving tank and the ESD was not activated automatically by the ESD system.
UCA2.19 Operator did not activate ESD manually when the ESD system had high pressure in the LNG receiving tank, and the ESD has not activated automatically.	N/A	UCA2.20 Operator did activate the ESD manually too late, when a High-high pressure in the LNG receiving tank, and ESD is not activated automatically by the ESD system.
UCA2.21 Operator did not activate the ESD manually when the ESD signal was received from receiving vessel, and ESD is not activated automatically by the ESD system.	N/A	UCA2.22 Operator did activate the ESD manually too late when the ESD signal was received from receiving vessel, and the ESD was not activated automatically by the ESD system.
UCA2.23 Operator did not activate the ESD manually when fire onboard was detected, and the ESD is not activated automatically by the ESD system.	N/A	UCA2.24 Operator did activate the ESD manually too late when fire onboard was detected, and the ESD was not activated automatically by the ESD system.

Table AB. 3 Control action3. Overriding

Not providing Cause hazard	Providing Causes Hazard	Timing error Too early/ late
N/A	UCA3.1 Operator did override to stop the ESD activation by the system when gas detection in cargo machinery space above 60% LEL. (H1)	N/A
N/A	UCA3.2 Operator did override to stop the ESD activation by the system when gas detection at bunkering manifold area above 60% LEL.	N/A
N/A	UCA3.3 Operator did override to stop the ESD activation by the system when the pressure at the bunkering manifold was high.	N/A
N/A	UCA3.4 Operator did override to stop the ESD activation by the system when the pressure in the vapour return line was high.	N/A
N/A	UCA3.5 Operator did override to stop the ESD activation by the system when the ESD valve motive power was lost.	N/A
N/A	UCA3.6 Operator did override to stop the ESD activation by system When the ERC was activated by default.	N/A
N/A	UCA3.7 Operator did override to stop the ESD activation by the system when High-high liquid level in an LNG receiving tank.	N/A
N/A	UCA3.8 Operator did override to stop the ESD activation by the system when High-high pressure in an LNG receiving tank.	N/A
N/A	UCA3.9 Operator did override to stop the ESD activation by the system when the ESD signal was received from receiving vessel.	N/A
N/A	UCA3.10 Operator did override to stop the ESD activation by the system when fire onboard was detected.	N/A

Table AB. 4 Control action 4. Activate ESD automatically

Not providing Cause hazard	Providing Causes Hazard	Timing error Too early/ late
UCA4.1 ESD logic controller did not activate the ESD when gas detection in cargo machinery space above 60% LEL. (H1)	UCA4.2 ESD logic controller activated the ESD when gas detection in cargo machinery was normal. (H1)	N/A
UCA4.3 ESD logic controller did not activate the ESD when gas detection at the bunkering manifold area above 60% LEL. (H2)	UCA4.4 ESD logic controller activated ESD when gas detection at the bunkering manifold area was normal. (H2)	N/A
UCA4.5 ESD logic controller did not activate the ESD when the pressure at the bunkering manifold was high. (H3)	UCA4.6 ESD logic controller activated the ESD when the pressure at the bunkering manifold was normal. (H3)	N/A
UCA4.7 ESD logic controller did not activate the ESD when the pressure in the vapour return line was high.	UCA4.8 ESD logic controller activated the ESD when the pressure in the vapour return line was normal.	N/A
UCA4.9 ESD logic controller did not activate the ESD when ESD valve motive power was lost.	N/A	N/A
UCA4.10 ESD logic controller did not activate the ESD When ERC was activated by default.	N/A	N/A
UCA4.11 ESD logic controller did not activate the ESD when a high-high liquid level was in the LNG receiving tank.	N/A	N/A
UCA4.12 ESD logic controller did not activate the ESD when High-high pressure in the LNG receiving tank.	N/A	N/A
UCA4.4.13 ESD logic controller did not activate the ESD when ESD signal was received from receiving vessel.	N/A	N/A
UCA4.14 ESD logic controller did not activate the ESD when fire onboard was detected.	N/A	N/A

Table AB. 5 Control action 5. Activate ESD manually

Not providing Cause hazard	Providing Causes Hazard	Timing error Too early/ late
UCA5.1 Site operator did not activate the ESD manually when gas detection in cargo machinery space above 60% LEL and ESD was not activated by other controllers. (H1)	UCA5.2 Site operator activated the ESD manually when gas detection in cargo machinery was normal. (H1)	UCA5.3 Site operator did activate the ESD manually too late when other controllers did not activate gas detection in cargo machinery space above 60% LEL and ESD. (H1)
UCA5.4 Site operator did not activate the ESD manually when gas detection at the bunkering manifold area above 60% LEL, and ESD was not activated by other controllers. (H2)	UCA5.5 Site operator activated the ESD manually when gas detection at the bunkering manifold area was normal. (H2)	UCA5.6 Site operator did activate the ESD manually too late when other controllers did not activate gas detection at the bunkering manifold area above 60% LEL and ESD. (H2)
UCA5.7 Site operator did not activate the ESD manually when the pressure at the bunkering manifold was high, and other controllers did not activate ESD. (H3)	UCA5.8 Site operator activated the ESD manually when the pressure at the bunkering manifold was normal. (H3)	UCA5.9 Site operator did activate the ESD manually too late when the pressure at the bunkering manifold was high, and other controllers did not activate ESD. (H3)
UCA5.10 Site operator did not activate the ESD manually when the pressure in the vapour return line was high, and other controllers did not activate ESD.	UCA5.11 Site operator activated the ESD manually when the pressure in the vapour return line was normal, and other controllers did not activate ESD.	UCA5.12 Site operator did activate the ESD manually too late when the pressure in the vapour return line was high, and other controllers did not activate ESD.
UCA5.13 Site operator did not activate the ESD manually when the ESD valve motive power was lost, and other controllers did not activate ESD.	N/A	UCA5.14 Operator did activate the ESD manually too late, the ESD valve motive power is lost, and other controllers did not activate ESD.
UCA5.15 Operator did not activate ESD manually when ERC was activated by default and ESD was not activated by other controllers.	N/A	UCA5.16 Operator did activate ESD manually too late, ERC was activated by default, and other controllers did not activate ESD.

UCA5.17 Site operator did not activate the ESD manually when Other controllers did not activate a high-high liquid level in the LNG receiving tank and ESD.	N/A	UCA5.18 Site operator did activate the ESD manually too late when Other controllers did not activate a high-high liquid level in the LNG receiving tank and ESD.
UCA5.19 Site operator did not activate ESD manually when Other controllers did not activate high-high pressure in LNG receiving tank and ESD.	N/A	UCA5.20 Site operator did activate the ESD manually too late when Other controllers did not activate high-high pressure in the LNG receiving tank and ESD.
UCA5.21 Site operator did not activate the ESD manually when the ESD signal was received from receiving vessel, and other controllers did not activate the ESD.	N/A	UCA5.22 Site operator did activate ESD manually too late when ESD signal was received from receiving vessel, and other controllers did not activate ESD.
UCA5.23 Site operator did not activate ESD manually when a fire on board was detected, and other controllers did not activate ESD.	N/A	UCA5.24 Site operator did activate ESD manually too late when a fire on board was detected, and other controllers did not activate ESD.