Doctoral Dissertation
Doctoral Program in Chemical Engineering (36^{th}cycle)

Evolving Process Maintenance through Human-Robot Teaming:
An Integrated System Performance Analysis

By

Shuo Yang

******

Supervisor(s):
Prof. Micaela Demichela, Supervisor

Doctoral Examination Committee:
Prof. A.B. , Referee, University of...
Prof. C.D, Referee, University of...
Prof. E.F, University of...
Prof. G.H, University of...
Prof. I.J, University of...

Politecnico di Torino
2024

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data. This research work was funded by the Chinese Scholarship Committee.

Shuo Yang
2024

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

I would like to dedicate this thesis to my loving parents

Acknowledgements

I will always remember the three years of Ph.D. study times I spent in Turin, Italy. I love this quiet and beautiful city. I love the warm-hearted people here. I love the architecture and culture here. It is not easy work for one to restart study life after a 10-year working period. But I really want to make a change. I really want to have an opportunity to have a deeper look at myself and what I want to devote myself to. In the end, I make it, with all the help from my supervisor, colleagues, and friends.

Firstly, I want to express my gratitude to my supervisor, Prof. Micaela Demichela. I still remember the scenario in which she welcomed me at the front door of Polito with a smile on the first day I came here. During my three years of Ph.D. study in Italy, she supported me academically and in life. I was worried about the adaptive period I needed to start a new stage in Italy. Well, her detergent working spirit and humor talk style let me adapt to the safer team pace quickly.

Special thanks to my friend Jie Geng. She introduced me to the Ph.D. opportunity and gave me company all the way through the Ph.D. times remotely in China. We did project applications together, wrote articles, and shared knowledge and ideas, and geographical distance cannot obstruct us.

I would like to give my thanks to my colleagues Carlos, Chidela, David, Gabriele, Huxiao, Morena, Joseph, and Francesca. I am grateful that I have been in a warm and creative team as safer.

Thanks to the technicians from Zhejiang Special Equipment Inspections Institute. They gave me great support with my case study investigation.

I want to extend my deep thanks to my family: my mother, Hong Ma; my father, Guilin Yang; my grandpa, Li Ma; and my partner in life, Xiaohang Zhou. They constantly supported me and pulled me outside the depression whenever I felt down.

Last but not least, I want to thank Polito. It is the end of my Ph.D., also a starting point for the future, thanks to teaching me that “what you are, takes you far.”

Abstract

Safety management is an essential subject in the process industry since accidents in this domain can have severe social, economic, and environmental consequences. Maintenance activities with functions for early fault detection are a critical part of the safety management system in the process industry.

Maintenance activities in the process industry often involve hazardous working conditions with high temperatures, low-quality air, and high physical demand. In the background of Industry 4.0, robot automation is introduced to maintenance operations to try to release human operators from tedious, repeated physical work and dangerous operations in maintenance activities. Human operators still need to be in the loop because of their ability to be more flexible and have a better strategic view. Therefore, human-robot collaborative teaming(HRT) operations will be a primary paradigm.

HRT operations involving multiple twisted elements are complex socio-technical systems for which previous methods that only focus on the technical or social part are not sufficient. These complex systems require a novel integrated framework able to consider human and organizational factors(HOFs), technical factors, as well as their interdependence.

Through data analysis of accident reports from eMARS, the most contributing organizational factors in HOFs and maintenance-related accidents were identified. Therefore, the Cognitive Reliability and Error Analysis Method(CREAM), which contains all these contributing organizational factors in the form of common performance conditions(CPCs), is selected to perform the HRA part in the integrated framework in an extended way.

After reviewing the literature on probability risk assessment, HRA, organizational factors, and complex system theory. The integrated risk-based performance assessment framework is proposed. This framework contains a qualitative phase

employing the top-down approach and a quantitative phase employing the bottom-up approach. The outputs of the first phase include a qualitative task analysis list supporting HRT team structure transitions analysis and task logical interdependence analysis. Then, the outputs of the second phase include quantitative risks and performance indicator values supporting system performance comparison and critical scenarios and factors analysis.

The proposed framework was demonstrated step-by-step in a case study of LPG spherical storage tank inspection in the full manual(FM) and HRT operations. The system performance of these two scenarios was compared. The results show the evolution in terms of the systemic performance from the FM to the HRT system. Also, the boundaries of critical factors were calculated. The quantitative results could provide better insights to support decision-making about HRT operation in practice. The framework validated is able to give guild to risk and performance assessment for complex socio-tech systems. Further work could be performed to extend the model with a reinforcement learning algorithm to optimize the robot paths and HRT schedule.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The development and implementation of robotics significantly impact safety management in the process industry, especially in maintenance activities, where the working conditions are extremely hazardous. Due to widely expanding digitization applications in Industry 4.0 and the human-centered paradigm in Industry 5.0, human-robot collaboration teaming(HRT) tends to be a primary operation scenario. Initially, this trend aimed to access hazardous locations that human operators can not or hardly get access to and improve productivity. In this research, the definition of HRT made by Lu et al. is adopted [2]:

"HRT studies how humans and robots work simultaneously in a shared space for a shared goal. While HRT involves accomplishing a shared purpose through joint efforts, it does not require humans and robots to share space through its flexible interaction between humans and robots, communication through remote control, or close collaboration."

HRT systems involve multiple interdependent elements: human operators, equipment, and robots are complex socio-technical systems. The literature review shows that no validated methodology could be utilized for the HRT system performance analysis [3]. Therefore, the framework for performance assessment of these systems is the main focus.

1.1 Background

Safety is significant for the process industry for its inherent attributions to involve a bulk amount of hazardous chemical materials and reactions, which have the potential to bring about accidents that would have serious consequences affecting people's lives, property, and the environment [4].

Maintenance activities are a critical part of process safety management. Firstly, the main function of maintenance is safety-related. According to European Standard [5], industrial maintenance is defined as:

"combination of all technical, administrative, and managerial actions during the life cycle of an item intended to retain or restore it to a state where it can perform the required function. Technical maintenance actions include observation and analyses of the item state (e.g., inspection, monitoring, testing, diagnosis, prognosis, etc.) and active maintenance actions (e.g., repair, refurbishment)".

The mission of maintenance is fault detection, maximizing the system availability and efficiency of the facility, ensuring its safe and proper functionality, and minimizing costs [6]. Secondly, accidents in the process industry were highly linked to maintenance activities. Based on data from the British Health and Safety Executive and database FACTS, about 30-40% of accidents were connected with maintenance activities, either during the maintenance period or due to faulty or no appreciable maintenance [7]. In a study of 183 major accidents in the process industries from 2000 to 2011, 44% were maintenance-related [8].

Traditionally, maintenance activities in the process industry require human operators to work in hazardous working conditions, including exposure to substandard air quality, dangerous materials, high noise, extreme temperatures, and intense physical demand [9] [10]. Therefore, the inspection robots designed to relieve human operators from hazardous environments have aroused great interest. As statistical data shows, there were more than 30,000 chemical industry robots in application in the year 2023 [11], and it is predicted that there will be around a 30% increase in revenue every year from 2023 to 2028.

According to the review about robotics applications in onshore oil and gas facilities, most of the robotic research has been devoted to developing in-pipe inspection robots (IPIRs) and tanks inspection robots [12]. A famous tank inspection robot, Neptune [13], could demonstrate the general features of the tank inspection robot.

The critical technologies of this kind of robot include a steering mechanism, a propelling mechanism, and detection technology. Tank inspection robots usually need to perform tasks on the curved surface. This requires the mechanism of moving stable on the curved surface without falling. The most common adhesion mechanisms are magnetism and vacuum suction. In addition, they must equip navigation systems to move along the setting paths automatically. Detection technology includes magnetic particle inspection (MPI) and ultrasonic. Most robots still need a tethered cable, power supply, and control signals as a function of a safety rope. According to the automation degrees, tank inspection robots can be categorized as manually operated, semi-autonomous, and fully autonomous. Whereas most tank inspection robots are remotely operated vehicles carrying appropriate sensors for detection and communication. Most of these robots still heavily depend on skilled operators for their operations [12].

1.2 Evolving From the Full Manual to HRT operations

Although robotic technology develops rapidly in the background of Industry 4.0, it is still not a fully automatic process in maintenance operations in the process industry. Human operators still need to be in the loop for their flexibility and strategic decision-making ability. Moreover, in the Industry 5.0 paradigm, robotics are designed to support rather than replace human operators. Therefore, HRT will be the primary operation scenario in the process maintenance activities. According to the workspace, working time, working goals, and existence of physical contact, the human-robot interactions can be divided into three subcategories, as Table 1.1 shows: human-robot coexistence, human-robot cooperation, and human-robot collaboration [14].

Table 1.1 Subcategories of human-robot interaction

Interaction types	Workspace	Work time	Work goals	Physical contact
Collaboration	same	same	same	yes
Cooperation	same	same	same	-
Coexistence	same	same	-	-

All three of these modes could be involved in the process maintenance operations. Therefore, the term HRT represents the human-robot collaborative teaming process. In HRT, robotic precision complements human flexibility and vice versa, enabling more efficient task delivery than either party could achieve alone[2].

As discussed in literature [15] [16], introducing robots in teams with humans is usually presented to reduce the operators' physical workload while increasing productivity. This benefit could be emphasized when applied in maintenance activities in the process industry, for it also reduces the time of human operators' direct contact with hazardous and dirty working conditions. However, this application will introduce new workflow and resource allocation changes. On the other hand, the new actors in the work environment in a cage-free form have been recognized to increase psychological stress to humans [17]. Decision-makers must consider these changes and their influences before robotics is implemented. A risk-based performance analysis could be used as a basis for decision-making because of its capacity to generate information about risks and performance uncertainties. However, there are several challenges in developing a framework for risk-based performance analysis for the HRT maintenance activities, which will be discussed in the next section.

1.3 Related Works and Challenges

The HRT process maintenance activities can be regarded as socio-technical systems. While in the risk assessment domain, Rasmussen summarized the evolution of risk reduction and safety management theories as a multiple-discipline merged process [18]: In the first stage, researchers focused on the technical aspects, the accident records analysis guided the studies to the human-machine interface problems, then entered the human error analysis area. In the second stage, research hotpots drifted into the work conditions for human workers management at the organizational level.

Along this line, the research boundary expanded again, including the safety regulation in law and government. After that, Rasmussen proposed that the complex systems risk analysis and management framework building requires a systems approach based on functional abstraction instead of structure decomposition to cope with the dynamic and highly integrated socio-technical systems.

The first challenge in developing an integrated risk and performance assessment framework for socio-technical systems is selecting the Human Reliability analysis (HRA) method. After more than 40 years of development, about 50 HRA methods have been developed, but some problems still impede the wide use of HRA in practice. Many HRA methods such as Technique for Human Error Rate Prediction (THERP) [19], Cognitive Reliability and Error Analysis Method (CREAM) [20], Connectionism Assessment of Human Reliability (CHAR) [21], Information, Decision, and Action in Crew Context (IDAC) [22], use performance shaping factors (PSFs), also called performance influencing factors (PIFs), common performance conditions (CPCs), and so on, to represent the aspects of the human-system influence, but there is still no standard PSFs set. literature criticizes these factors as a lack of validation and data support [23].

The second challenge concerns the representation and analysis of elements' interdependence. HRT maintenance operations in the process industry involving multi-interacted elements are complex socio-technical systems.

The complex system theory summarizes the main characteristics of these systems: interdependency, non-linearity, and emergence. Sun et al. [24]identified potential safety risks regarding the construction industry's physical, attention costs, and psychological impacts. Their research offered a comprehensive analysis of robots' negative impact on workers' safety and health. Nevertheless, it focused solely on the safety dimension of system performance and yielded qualitative findings. Borges et al. [15]proposed an integrated framework based on a system dynamics model to predict the workstation performance after implementing the HRT system while only focusing on the ergonomic aspects.

Reviews Also highlighted the main shortcomings that limit the practical applicability of novel risk assessment methods to HRT: fragmentation, complexity, and lack of validation [3]. In addition, the need to consider the factors that could affect the safety of operators and the efficiency of the operations in the long term was identified [17] as the need for integrating workers' personal aspects into the hazard analysis in the form of more human-centered specific guidelines to suit the HRT context better [25]. Therefore, the main gaps in current performance assessment methods regarding HRT complex systems include the following:

1. They usually ignore the HOFs or oversimplify the human error mechanism using static point value to represent the human error probability (HEP) or model the human error as a Poisson Function.

2. Many researchers focus on the HRA or technical, functional risk assessment, but the integrated framework considering the elements' interdependence is lacking.

3. In addition, there is a need to investigate the detailed data about the system performance changes during the FM scenarios evolving to the HRT scenarios.

Advancement with respect to the state of the art. The proposed framework differentiates from the reference literature(in Chapter3) to fill the gaps mentioned in last paragraph. Overall, it is built upon the complex system theory in a holistic manner. Specifically, for three main aspects: First, the CREAM method is extended with the Dempster-Shafer theory to reduce subjective bias, and fatigue-recovery functions are employed to achieve more realistically dynamic attributes representing human behavior. Secondly, the integrated dynamic decision analysis (IDDA) method is employed to represent the subtasks' interdependence in a logically constrained way. Meanwhile, the agent-based modeling and simulation (ABMS) method models the HRT communications by message change mechanism. Thirdly, the proposed framework is tested in a case study to provide validation.

1.4 Logical Structure and Research Questions

Based on the complex systems theory as discussed in Section 1.3 and trying to fill the gaps in the literature, six studies, as listed in Table 1.2, were conducted to explore the main aspects of performance assessment for complex socio-technical systems, specifically with a holistic approach to deeply understanding how to use reliable data of elements and their interdependence to forecast the performance of HRT systems.

Table 1.2 Research Publications

Title	Journal	Time
Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations	Safety Science	2024[26]
A data-driven narratives skeleton pattern recognition from accident reports dataset for human-and-organizational-factors analysis	Journal of Loss Prevention in the Process Industries	2023[1]
Analysis of human and organization factors related accident reports based on natural language processing	Chemical Engineering Transactions	2022[27]
Contributions and Consequences Coming from Human and Organizational Factors to the Accidents	Chemical Engineering Transactions	2022[28]
Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis	Advanced Engineering Informatics	under review
A Data-driven Bayesian Network of Management and Organizational Factors for Human Reliability Analysis in the Process Industry	Heliyon	under review[29]

Under the two challenges mentioned in Section 1.3, and the gaps in the literature. Formally, four main research questions derived from the necessity to comprehend the abstract construction of *integrate* framework include the following:

1. RQ1: What are the most critical HOFs contributing to accidents in the process industry? Accident reports are informative data sources from which to learn from the past. To understand the main elements of the complex social-technical system better, especially HOFs. In Chapter 2, the eMARS dataset is selected to identify the most contributing HOFs and their interdependent mechanism. Two related studies have been published in Chemical Engineering Transactions, articles entitled "Contributions and Consequences Coming from Human and Organizational Factors to the Accidents"(2022) [28] and "Analysis of human and organization factors related accident reports based on natural language processing"(2022) [27]. One paper published in the Journal of Loss Prevention, entitled "A data-driven narratives skeleton pattern recognition from accident reports dataset for human-and-organizational-factors analysis"(2023) [1].

2. RQ2: What is the integrated system performance evaluation framework? In Chapter 3, a detailed literature review on systematic risk assessment and human reliability analysis methods is carried out. This contributes to the theoretical building blocks for the integrated framework. In Chapter 4, the comprehensive top-down and bottom-up framework, considering the organizational, human, equipment, and robot elements and their interdependence, is proposed and explained step by step. Also, In Chapter 5, the proposed framework was applied in a case study about pressurized tank inspection activities. This research question was addressed first in a study published in the journal paper entitled "Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations."(2024) [26]

3. RQ3: What are the critical differences between HRT and FM system performance? The case study results show the answer to this research question in qualitative and quantitative ways in Chapter 6. They are also included in the paper "Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations"(2024) [26] and the paper under review entitled "Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis."

4. RQ4: What are the primary parameters influencing the performance of HRT and FM systems? In Chapter 6, a sensitivity analysis was performed to test the proposed method's logical consistency and explore the strength of the parameters' influence on the overall system performance. This part is also included in the paper "Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis" in the journal Advanced Engineering Informatics, which is under review.

The rest of this thesis is organized as follows: Chapter 2 performs HOFs contributed Maintenance-related Accidents Data Analysis: To understand the HOFs factors better, this research collected accident reports data from the eMARS dataset. Based on the Natural Language Processing (NLP) techniques, The "4Ws" model was proposed for the HOFs-contributed accident analysis. The characteristics of HOFs contributing to the accident were analyzed by comparing "4Ws" information with other accidents. Chapter 3 gives the theoretical basis: The state-of-the-art and research gaps in systemic risk assessment. The overview of the development of HRA methods and their conceptual framework. Summary of the result of organizational factors analysis for HRA. The basic theory is the foundation of a complex system. Chapter 4 presents the integrated framework: The integrated framework provides both qualitative and quantitative system performance evaluation approaches considering HOFs. Chapter 5 applies the proposed framework to a case study in the pressurized vessel inspection process: Step-by-step applications of the framework are conducted (in the gas and oil industry). Chapter 6 demonstrates the Results: the traditional FM and HRT scenarios were compared to see the difference between the two systems' performance and better understand the HRT process to support future decision-making. Chapter 7 discusses the result compared to previous research and the limitations. Chapter 8 concludes the thesis and gives further research directions.

Chapter 2

Learning from the past to understand HOFs

Unlike the ever-changing technical factors, the HOFs, especially the organizational factors, are relevantly stable, which makes it more possible to learn from the past. Therefore, to better explore the HOFs' contributions to process accidents and their inter-dependency, the accident reports from the eMARS dataset were investigated in this Chapter. Two works have been published in Chemical Engineering Transactions, entitled "Analysis of Human and Organization Factors Related Accident Reports Based on Natural Language Processing" [27] and "Contributions and Consequences Coming from Human and Organizational Factors to the Accidents" [28], also one in the journal Process Safety and Loss Prevention entitled "A Data-driven Narratives Skeleton Pattern Recognition from Accident Reports Dataset for Human and Organizational Factors Analysis" [1], another one in the journal Heliyon entitled"A Data-Driven Bayesian Network of Management and Organizational Factors for Human Reliability Analysis in the Process Industry." [29]

2.1 Data Source

A significant obstacle to HOFs-related research is the lack of relevant data. However, accident reports emerge as a valuable data source, given that they are reliable information derived from actual incidents and are validated by authoritative entities.

The eMARS dataset[1] encompasses an extensive collection of documents on chemical accidents and near-miss incidents reported to the Major Accident Hazards Bureau of the European Commission's Joint Research Centre by member states of the EU, EEA, OECD, and UNECE countries (in alignment with the TEIA Convention).

Within the eMARS database containing 73 data columns, a subset of ten columns were selected in this HOFs contribution analysis. These columns include Accident ID, Human, Organizational Causative Factor Type, Human On -site Quantity/Effect, Environmental On-site Quantity/Effect, Cost On-site Quantity/Effect, Disruption Off-site Quantity/Effect, accident description, causes of the accident, and lesson learned. This research focuses on the analysis of cases pertaining to HOFs. Initially, from the database, 1128 cases were excluded for not containing/identifying 'Human' or 'Organizational Causative Factor Type' factors, culminating in the selection of 531 cases directly relevant to the HOFs. The preliminary statistical overview of the database is shown in Table 2.1 [27].

Table 2.1 Statistic overview of the database[1]

Total	Human error cases	Organizational factor cases	Redundant cases	No/too simple cause description cases	Other cases	Final HOFs cases
1128	209	464	142	42	639	489

2.2 Methods

For the factor influencing mechanism analysis based on categorical data, the Chi-square test is employed. For the knowledge extract from accident reports data, the "4W"s model is proposed.

2.2.1 Chi-square test

The Chi-square test, a pivotal statistical tool, examines the association between categorical variables that are not inherently ordered. Its primary objective is to ascertain the presence of a statistically significant discrepancy between observed frequencies in the data and the frequencies that would be expected under a specific hypothesis. This test operates on the principle that a higher Chi-square value indicates a more

[1]accessible at https://emars.jrc.ec.europa.eu/en/emars/accident/search.Retrieved November 17, 2022

pronounced divergence between the observed and expected values. Conversely, a lower value signifies minimal deviation. A Chi-square value of zero denotes perfect alignment, suggesting that the observed data perfectly match the expected theoretical frequencies. Central to the Chi-square analysis is the testing of the null hypothesis, which posits that there is no significant difference between the observed outcomes and those anticipated. The null hypothesis serves as a benchmark for determining whether the observed variations are due to chance or reflect a genuine association between the variables under investigation. For the data analysis in this study, IBM SPSS Statistics 24 was employed, offering a robust platform for executing the requisite statistical evaluations [30].

The investigation into the contributions of human and organizational factors (HOFs) to process accidents employed the categorical data analysis method. This analytical approach is characterized by its focus on classified response variables, which are mutually exclusive and can be either ordered or unordered [31]. The transformation of categorical data pertinent to this study is detailed in Table 2.2 and Table 2.3. This methodological choice enables the systematic examination of the impact of discrete, categorical variables on the incidence and characteristics of process accidents, thereby facilitating a nuanced understanding of the role played by HOFs in such events.

Table 2.2 Data transformation for cause factors

Cause factors	Integrated categorical variables	
Human Error Factor	①operator error	1
	②malicious intervention	2
	③willful disobedience or failure to carry out duties	3
	④operator health (includes ailments, intoxication, death, etc.)	4
	⑤failure to carry out duties not identified	5
	⑥not know	6
Organizational Factor	①design of plant/equipment/system	1
	②installation/construction	2
	③process analysis	3
	④maintenance/ testing/inspecting	4
	⑤training/instruction	5
	⑥supervision/staffing	6
	⑦user-unfriendliness	7
	⑧management attitude problem	8
	⑨organized procedures/management organization inadequate	9
	⓪not known/not applicable/empty	0

Table 2.3 Data transformation for consequence category

Consequence category	Integrated categorical variables	
Human On or Off-Site Effect	At risk	1
	Injury	2
	Fatalities	3
	not known / not applicable	4
Environmental On or Off-site Effect	Freshwater Pollution	1
	Inland Pollution	2
	Offshore Pollution	3
	Atmosphere Pollution	4
	not known / not applicable	5
Cost On or Off-Site Effect	material losses	1
	response, cleanup, restoration costs	2
	fine and legal costs	3
	Production loss/ System Interruption	4
	Profit Failure	5
	not known / not applicable	6
Social Effect	Infrastructure influence (telecommunication, roads, railways, waterways, air transport, etc.)	1
	nearby factories, offices, small shops	2
	schools, hospitals, institutions	3
	nearby residences, hotels	4
	Other places of public assembly	5

2.2.2 Natural Language Processing Techniques

The "4Ws" information structure framework has been devised as a foundational tool for the analysis of accident reports influenced by Human and Organizational Factors (HOFs), as depicted in Fig.2.1. This framework focuses on capturing essential HOFs-related information, including when accidents occurred, the equipment and locations involved, the actors (individuals or groups), and the specific HOFs implicated. Accident reports are conceptualized as narrative stories, and this research aims to distill the core elements.

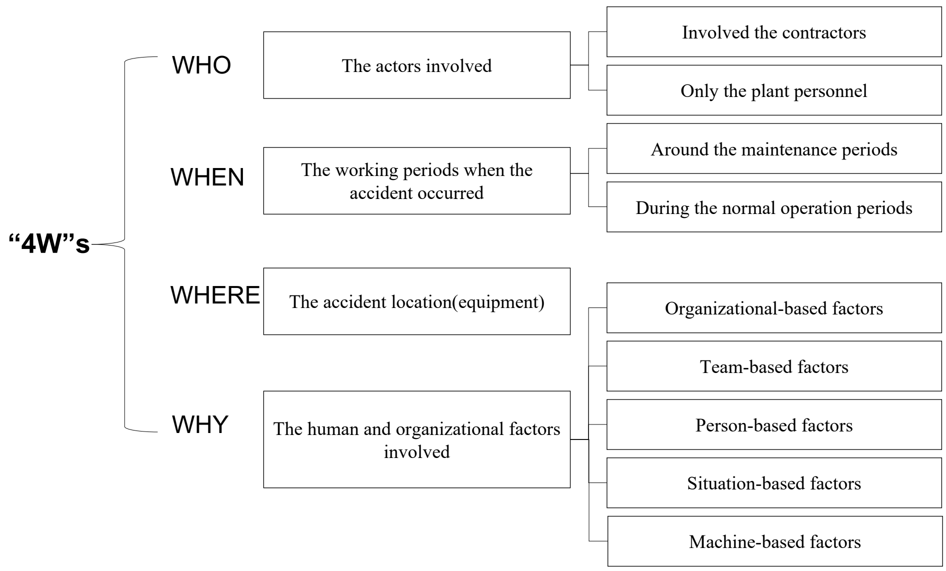


Fig. 2.1 The "4Ws" model

The traditional "5Ws and 1H" framework, prevalent in communication and journalism, outlines Who, What, When, Where, Why, and How. This structure is designed to orient readers to their communal environments, facilitating a connection beyond their immediate sensory perceptions and ensuring the delivery of information that satisfies their needs [32]. In the context of HOFs-focused risk communication, the primary audience comprises frontline operators and decision-makers. Mallam et al. [33] highlight that sharp-end operators exhibit a heightened interest in human factors-related issues over technical details, as these are more relatable to their professional identities and cultures.

The "4Ws" model for HOFs information adopts a simplified narrative frame—"Who, When, Where, Why"—by mining relevant details from accident records. This research posits two empirical hypotheses to address unique aspects of HOFs-related

Table 2.4 Tokens in spaCy[1]

TEXT	LEMMA	POS	TAG	DEP	SHAPE	ALPHA	STOP
the	the	DET	DT	det	X	TRUE	TRUE
fatigue	fatigue	NOUN	NN	subj	xxxx	TRUE	FALSE
of	of	ADP	IN	prep	xx	TRUE	TRUE
a	a	DET	DT	det	x	TRUE	TRUE
valve	valve	NOUN	NN	obj	xxxx	TRUE	FALSE
caused	cause	VERB	VBD	ROOT	xxxx TRUE	FALSE	
the	the	DET	DT	det	xxx	TRUE	TURE
block	block	NOUN	NN	obj	xxxx	TRUE	FALSE
of	of	ADP	IN	prep	xx	TRUE	TRUE
chlorine	chlorine	NOUN	NN	obj	xxxx	TRUE	FALSE
.	.	PUNCT	.	punct	.	FALSE	FALSE

accidents: first, the special emphasis on contractors due to their critical interface with owners for process safety [34]; second, the identification of maintenance, emergency, and control room operations as high-stakes areas where human intervention significantly increases the risk of human error [35]. Consequently, the "Who" component examines the involvement of either contracted or internal personnel in accidents. The "When" aspect focuses on whether incidents occurred during maintenance activities, while "Where" concerns the specific equipment or locations of the accidents.

Keywords were extracted utilizing the spaCy package [36], a sophisticated natural language processing (NLP) library within the Python programming ecosystem. The initial phase of pre-processing involved tokenization, a fundamental NLP task where raw text is segmented into 'docs' through spaCy's integrated pipelines. Tokenization entails the decomposition of text into sentences and words, exemplified by transforming the sentence "The fatigue of a valve caused the block of chlorine." into a structured 'doc' form, as depicted in Table 2.4. Within these 'docs,' individual tokens can be identified, manipulated, and subjected to further analysis.

The Named-entity Recognition (NER) process is utilized for the crucial task of keyword extraction. This involves the refinement of a pre-trained model, a hallmark of spaCy's functionality, which boasts a wide array of built-in models. This process facilitates the joint training of both newly incorporated classifier layers and the foundational layers of the base model. The extracted keywords play a pivotal role in supporting and enriching the analysis. The "When" and "Why" information have limited categories and some structured patterns. Therefore, the rule-based matching method was selected. While the "Where" information has unlimited examples, the fine-tuning model method was selected, as shown in Fig.2.2.

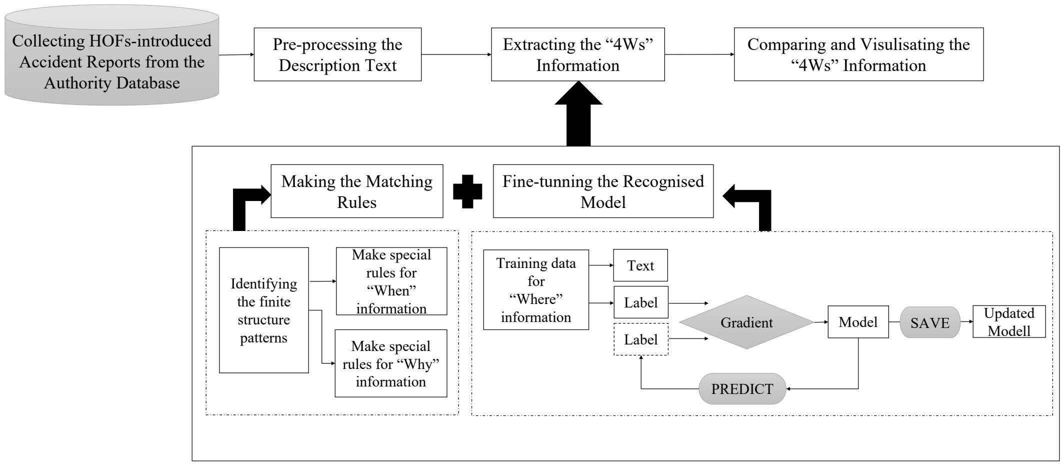


Fig. 2.2 The matching approaches selection

2.3 Results

2.3.1 The impact of HOFs on the recorded accident consequences

Fig. 2.3 and Fig. 2.4 present a detailed analysis of the contributions of different types of Human and Organizational Factors (HOFs) to the documented incidents. The category data were transformed according to Table 2.2. Notably, human errors were implicated in 40% of the recorded accidents, underscoring the significant impact of human causative factors. Furthermore, organizational causative factors contributed to a staggering 92.97% of the recorded accidents. Among these, a specific subset of organizational factors denoted as "⑨①⑤③④," accounted for 76.83% of the incidents. This data highlights the predominant role of organizational dynamics in the genesis of accidents, suggesting that interventions targeting these factors could yield substantial improvements in safety outcomes.

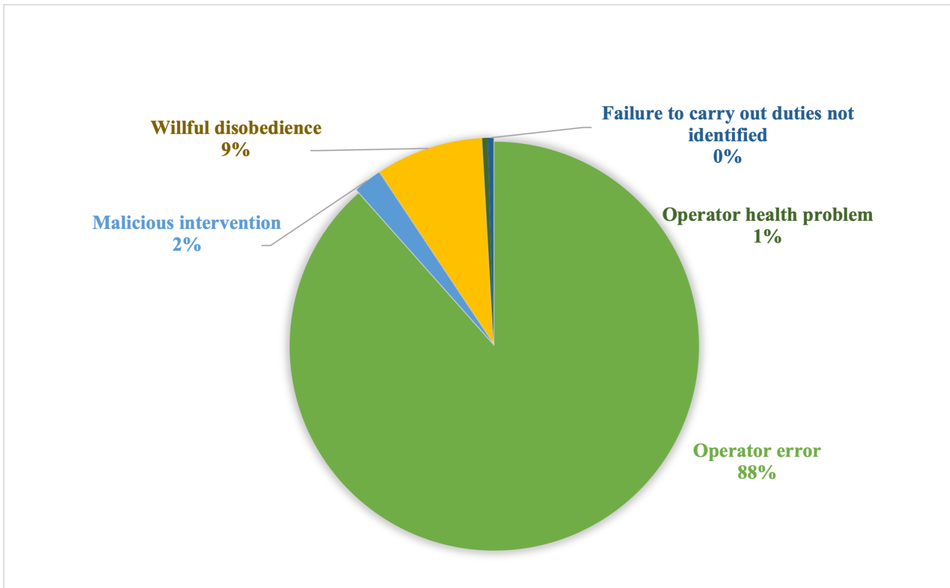


Fig. 2.3 Human error factors distribution

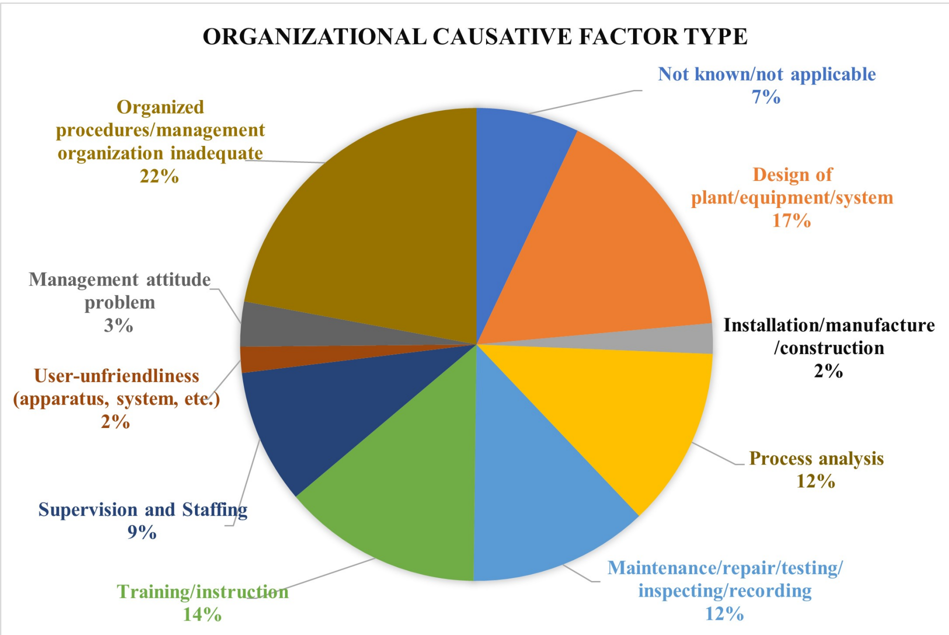


Fig. 2.4 Organizational factors distribution

Human errors factor contributions

To delve deeper into the pivotal influences of Human Errors on accident consequences, this study explored the "Human On-site Effect" and the "Cost Off-site Effect." Fig. 2.5 elucidates that the human error category of "①-operator error" was a contributing factor in 190 accidents, accounting for 88.37% of incidents when excluding cases categorized as 'Not known/not applicable.' Following this, "③-willful disobedience" was identified as contributing to 9% of the incidents, with the remaining 3.26% attributed to other causes. Furthermore, Figure 2.6 delineates the profound impacts of operator errors on the human on-site effect. Specifically, the outcomes associated with "injury (37.03%)", "injury and fatalities (28.40%)", and "fatalities (24.69%)" are emphasized, underscoring the severe consequences that can arise from such errors. Fig. 2.7 and Fig. 2.8 show the contribution of human error to the "Cost Off-site Effect."

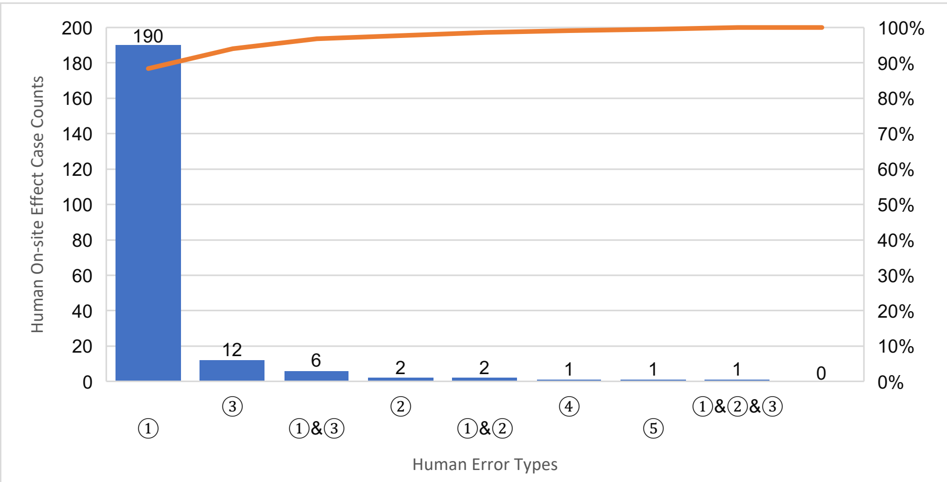


Fig. 2.5 Human errors contribution to human on-site effect

Organizational Factor contributions

This study extends its analytical depth to assess the paramount impact of Organizational Factors on accident consequences, specifically examining the "Human On-site Effect," "Environmental Off-site Effect," and "Cost Off-site Effect." Fig. 2.9 reveals that organizational causative factors encapsulated by the symbols "①③④⑤⑨" were implicated in 75.11% of incidents affecting humans on-site. These organizational factors' major influences are further dissected, with the consequences

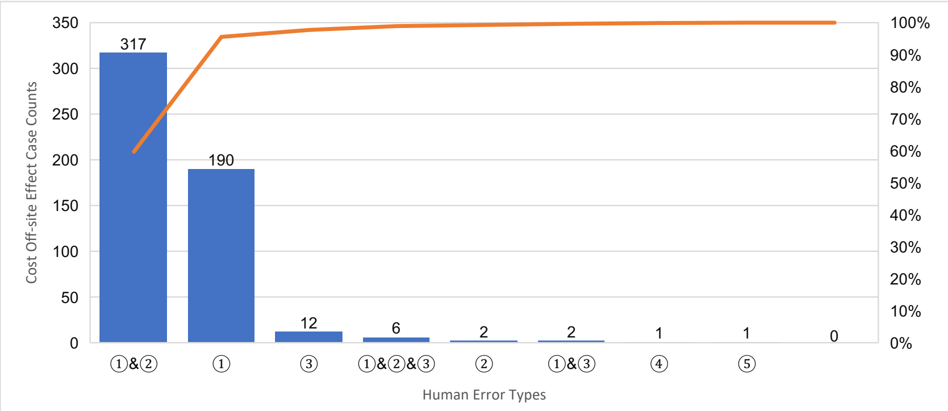


Fig. 2.6 Operator errors contribution to human on-site effect

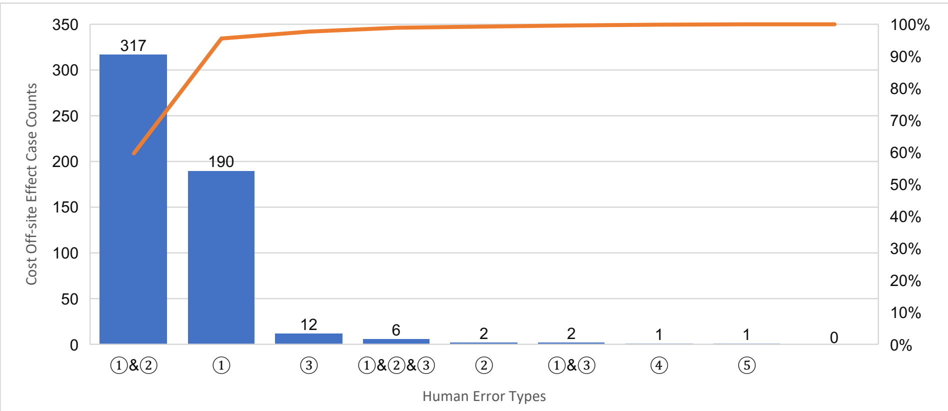


Fig. 2.7 Human errors contribution to cost off-site effect

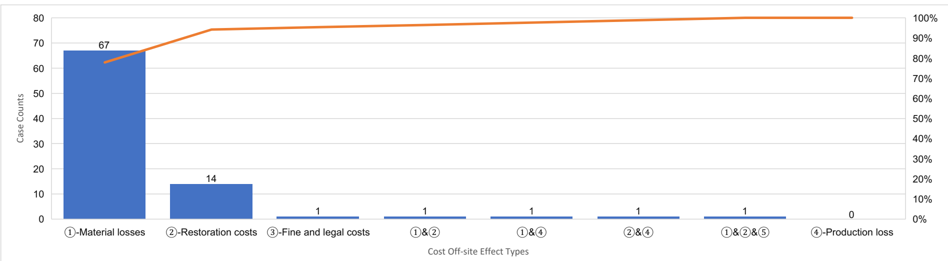


Fig. 2.8 Operator errors contribution to cost off-site effect

such as "injury" constituting 33.17% of all human on-site effects attributed to these factors, followed by "at-risk" conditions (28.65%), and "injury and fatalities" (22.71%). Moreover, among these organizational factors, "⑨-organized procedures/management organization inadequate" emerged as the most significant, contributing to 27.4% of accidents. This was followed by "①-design of plant/equipment/system" (22.44%), "④-maintenance/inspection" (17.43%), "⑤-training/instruction" (16.45%), and "③-process analysis" (16.29%). This granular analysis highlights these factors' critical roles in precipitating accidents, indicating areas where organizational improvements could mitigate such incidents.

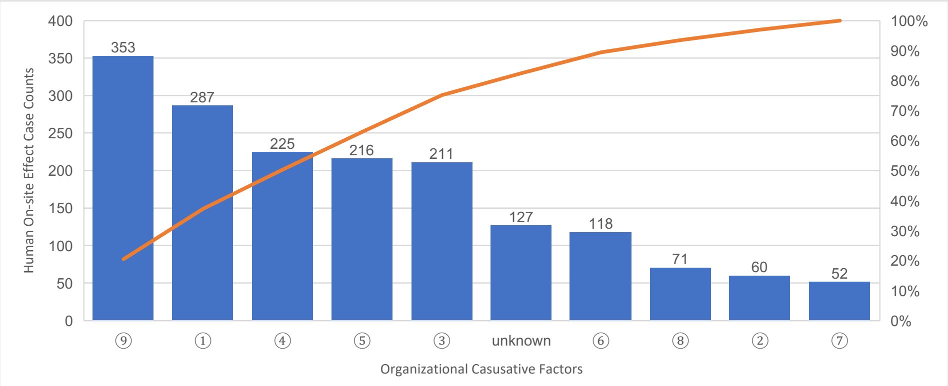


Fig. 2.9 Organizational factor contribution to human on-site effect

Fig. 2.10 delineates the significant contribution of Organizational Factors, represented as "①③④⑤⑨," to the environmental off-site effect, accounting for 76.95% of its impact. These factors' predominant influence is further elucidated, with "atmosphere pollution" emerging as a critical consequence, constituting 98.34% of all off-site environmental effects attributed to these organizational factors.

In particular, "⑨-organized procedures/management organization inadequate" was identified as the most significant contributor to atmospheric pollution, responsible for 27.99% of such incidents. It was followed by "①-design of plant/equipment/system" (23.05%), "⑤-training/instruction" (17.06%), "③-process analysis" (16.15%), and "④-maintenance/inspection" (15.76%). This detailed analysis underscores the pivotal role of specific organizational factors in exacerbating environmental impacts, particularly atmospheric pollution, and suggests targeted areas for intervention to mitigate these effects.

Fig. 2.11 illustrates the substantial role of Organizational Factors, denoted as "①②③④⑤," in accounting for 64.82% of the cost associated with off-site effects.

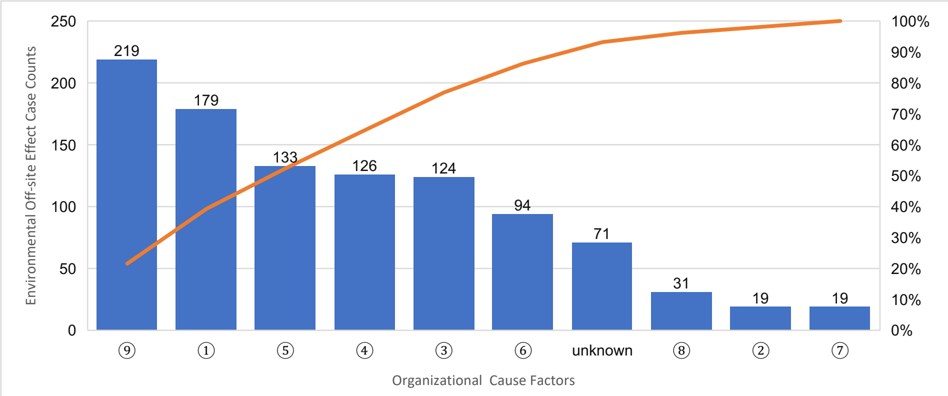


Fig. 2.10 Organizational factor contribution to environment off-site effect

Notably, "material losses" emerge as the predominant consequence, representing 72.86% of all cost off-site effects attributed to the identified organizational factors. Among these, the factor "①-design of plant/equipment/system" is pinpointed as the leading contributor to material losses, with a 32.41% share. This is followed by "②-Installation/construction" (22.07%), "③-process analysis" (17.93%), "④-maintenance/inspecting" (14.48%), and "⑤-training/instruction" (13.10%). This detailed breakdown highlights the critical impact of specific organizational factors on financial repercussions stemming from accidents, particularly in terms of material losses. It underscores the need for targeted improvements in organizational practices to mitigate these costly effects.

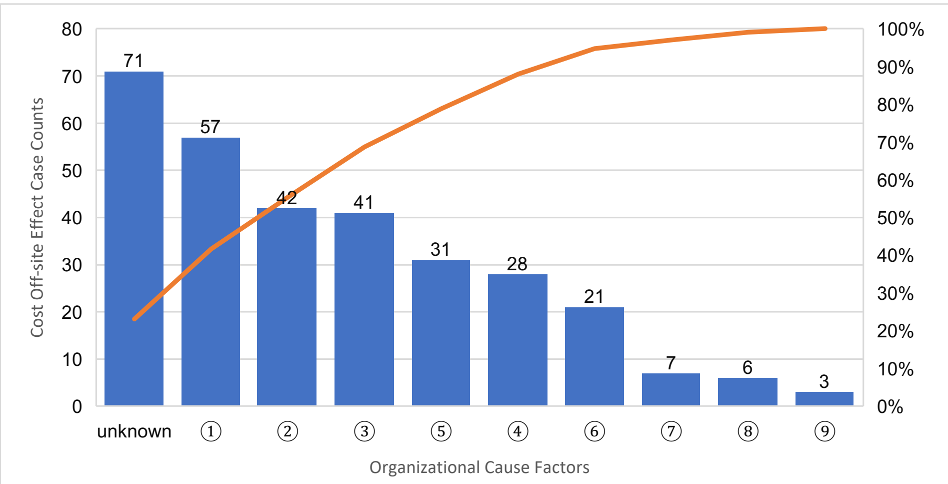


Fig. 2.11 Organizational factor contribution to cost off-site effect

2.3.2 “4Ws” information from accident reports

The analysis reveals a notable discrepancy in the incidence of Human and Organizational Factors (HOFs)-related cases during maintenance operations compared to other scenarios. Specifically, HOFs cases were nearly three times more likely to occur during maintenance, with approximately 20% of such incidents occurring in this context, in stark contrast to only 7% of other cases. This significant difference underscores the heightened vulnerability to human errors during maintenance activities and highlights the frequency of accidents attributed to insufficient procedures and instructions for maintenance tasks. Thus, decision-makers must prioritize maintenance periods within their risk mitigation strategies and safety resource allocation.

In a separate analysis, Single et al. [37] employed a rule-based Natural Language Processing (NLP) technique to categorize incidents within the eMARS database. From an examination of 889 cases, 77 distinct locations were pinpointed, accounting for merely 8.7% of the cases. This research extends beyond, analyzing 1128 cases and identifying 1728 instances of equipment involvement, indicating the involvement of multiple equipment pieces in certain accidents. Reactors were the most common site for HOFs-related incidents, followed by valves, tanks, and pipes, indicating specific areas of risk concentration.

The study further observes that contract operators were involved in 44 cases, with a significant portion (37 cases) engaged in maintenance tasks such as hot work, cleaning, repair, and replacement, besides five instances of transport services. This points to the critical need for stringent oversight of contract labor, especially during maintenance operations.

Regarding the classification of Precipitating Influencing Factors (PIFs), after distilling HOFs information, the analysis found 48 cases with three identified factors, 91 with two, 146 with one, and 167 with none. Organizational factors dominated the identified causes, constituting over two-thirds of the factors, with "procedures" (24%), "maintenance" (22%), "design" (18%), and "training" (10%) being the most prevalent. This distribution underscores the significant role of organizational elements in accident causation, suggesting a targeted focus on enhancing safety protocols and training initiatives.

2.3.3 Maintenance-related sub dataset analysis

There were 107 cases related to maintenance that had clear, reasoned investigation results. Fig. 2.12 shows these cases' most contributing equipment and process factors. In contrast, Fig. 2.13 shows the most contributing organizational factors. The first two equipment and process factors were corrosion/fatigue and component/machinery malfunction. They all contribute to over 20 percent of the maintenance-related cases.

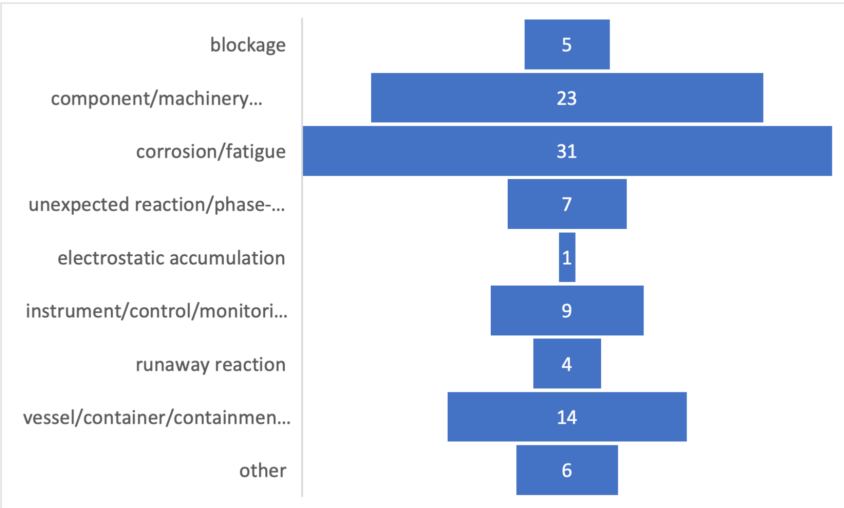


Fig. 2.12 Most contribute equipment and process factors in maintenance-related cases

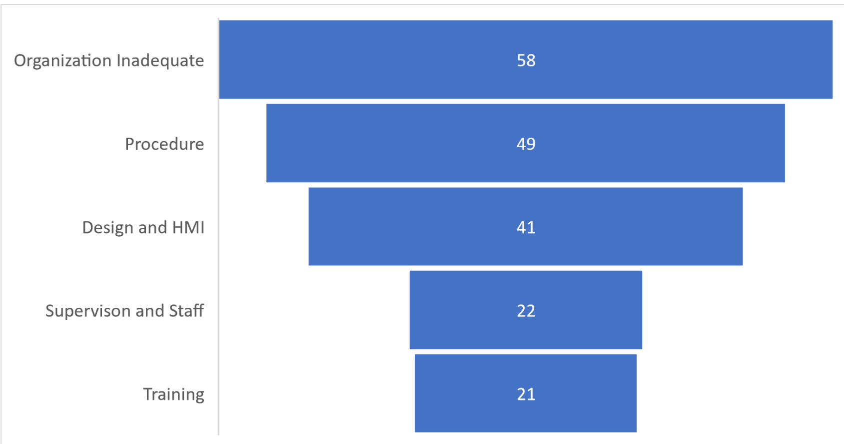


Fig. 2.13 Most contribute organizational factors in maintenance-related cases

The five most contributing organizational factors were inadequate organization, procedure, design, supervision, and training.

2.4 Bayesian Network of Organizational Factors

Organizational factors, called Management Delivery Systems, are usually the root cause of industrial accidents. They caught the attention of accident investigation researchers in the accident causal model just after the advent of industrial manufacturing in the 19th century, responding to the need for accident investigation practice. Heinrich (1941) proposed the falling domino stone accident causation model, which broke the mindset of always blaming the front-line operator and gave the concept of underlying or indirect causes. Going further in this direction, Reason [38] came up with the famous "epidemiological" model, the Swiss Cheese Accident Model, to metaphor the trajectory of how hazards go through "holes" of five levels of safety barriers. This model and tripod beta accident investigation method are applied, which lists 11 General Failure Types. Based on 300 naval aviation accident data, Shappell and Wiegmann [39] extend the Swiss cheese accident model framework by giving more specific definitions of the" barriers" and "holes," developing the HFACS model. The HFACS model divides organizational influence into resource management, organizational climate, and operational process. The organizational factors definition is lacking and not standard; the definition is collected from literature, as A.1 shows.

Further coded the above-mentioned eMARS data, a data-driven Bayesian Network (BN) can be learned and tested using the "bnlearn" package in R based on the accident reports dataset and visualized with GeNle software. "bnlearn" provides various machine-learning algorithms for BN structure construction and parameter estimate. GeNle is used to demonstrate BN visualization and calculate the impact strength and conditional probabilities range.

2.4.1 Structure learning

The three main categories of BN structure learning algorithms are constraint-based (CB), score-and-search (SS), and the hybrid method. The CB method is mainly based on statistical tests to identify a set of link constraints for the graph and then find the best DAG that satisfies the constraints to decompose the topological model construction.

The algorithm PC (named after its authors, Peter and Clark) is the most widely used CB method. This algorithm starts with a complete, undirected graph and deletes recursively edges based on conditional independence decisions [40]. Colombo et al. [41] proposed stable-PC, the algorithm aims to query all the neighbors of each node and fix these neighbors while conducting conditional independence tests (CI tests) at each level (based on the size of the conditioning sets) of the PC algorithm. The SS approach comprises the search algorithm and the definition of a score metric. The SS method views a structure learning process as a model-chosen problem, firstly finding all the possible network structures, utilizing a scoring function to assess how well the BN fits the data, and then using search algorithm searches over the space of DAGs to find the structure with the highest score[42]. The hybrid method combines CB and SS to utilize the strength of the two methods. Firstly, the variable sequence is obtained through the CI tests, which is then used as the input of the SS method to learn the final structure.

This study employs the hybrid method max-min hill climbing (MMHC) method to build the BN structure. This research codes each organizational factor state as a binary value 0 or 1 for "Success" or "Fail," so the node of BN has the Binomial distribution, which can be viewed as a particular form of the multinomial distribution. Based on this assumption, the stable-PC algorithm with a Chi-square as a CI test method is chosen firstly, with a p-value of 0.05, to reconstruct a Bayesian network's skeleton and then perform a Bayesian Dirichlet score-based hill-climbing search to orient the edges.

Twenty-three factors were analyzed using the bootstrap samples method to get a robust structure. The bootstrap samples were adopted to resample the data 5000 times, and learning one structure from each sample, then checking the frequency of one arc occurrence rate, then using the arcs with relevant high occurrence rates higher than 0.5 to build a robust network structure. The result of the BN is shown in Fig.2.14, based on which the organizational factors and human error-related parts are kept, then the final structure of BN is obtained, as Fig.2.15 shows.

2.4.2 Parameters learning

Estimate the parameters of the node using the Bayesian score, and then the conditional probability table (CPT) is gained. The example of CPT is shown in Table 2.5.

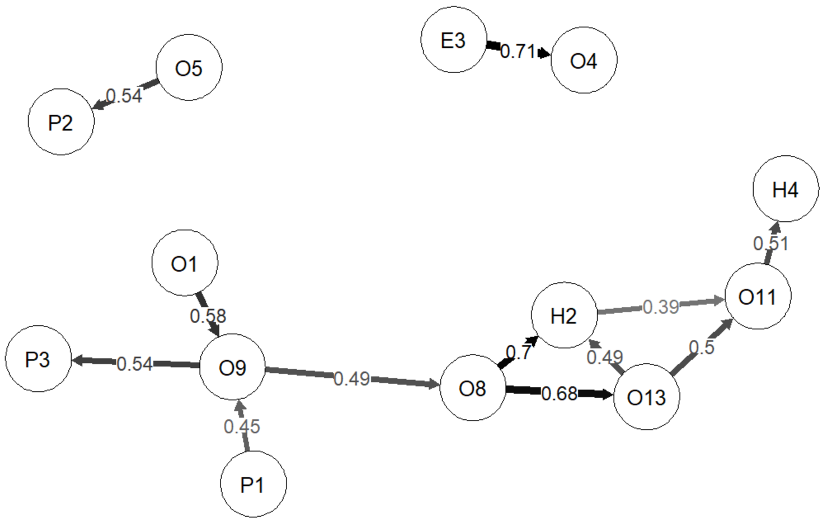


Fig. 2.14 Bootstrap MMHC structure learning (score= “BDe”, test = "x2", a=0.05)

Table 2.5 The conditional probability of variable Procedure

	Training			
	Success		Fail	
	Procedure		Procedure	
Operator Error	Success	Fail	Success	Fail
Success	0.8994647	0.6379310	0.5551724	0.3976109
Fail	0.1005353	0.3620690	0.4448276	0.6023891

2.4.3 BN validate and visualization

Fifty runs of 10-fold cross-validation were performed to validate the model learning strategy and measure the predictive accuracy for all the other variables. The mean of results is a classification error of ≈ 0.12.

The BN model comprises five dependent organizational variables: Design, Process Analysis, Procedures, Training, and Supervision. Procedures and training are two directly influential factors in Operator Error. Supervision is the directly influential factor in Willful Violations. The GeNle Academic Software is applied to visualize the Bayesian network and analyze the sensitivity. Since the data sample referred to an accident, the resulting probability of human error is higher than the literature reference values. The SPAR-H nominal value of the human error of action is 0.001. At the same time, our model gives back a value of 0.19. According to the accident dataset, the operator error and willful violation ratio is 9:1. Therefore, the operator error probability can be updated with set virtual evidence, such as the prior failure rate of 0.0009 for operator error and 0.0001 for willful violations rate. After the update, the whole BN is shown in Fig. 2.15.

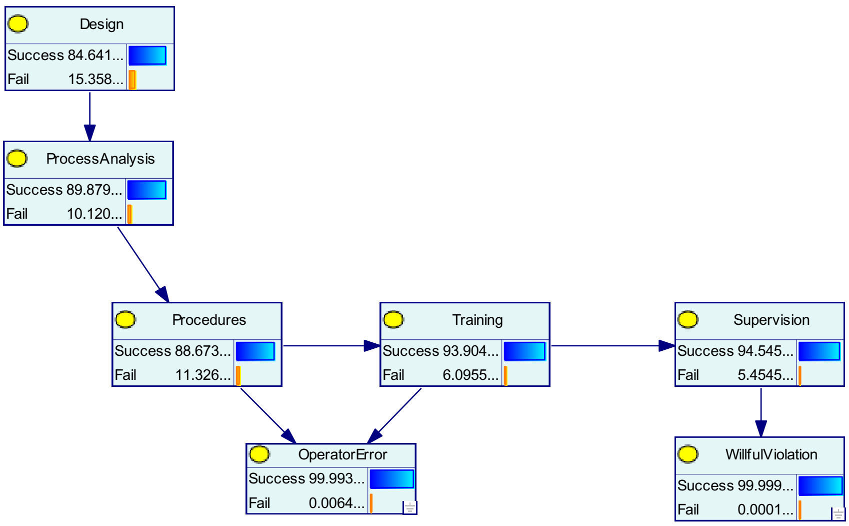


Fig. 2.15 BN structure of organizational factors to human error

The results of the sensitivity analysis of the Bayesian network are shown in Fig.2.16, where the darker color means the variables are more sensitive. The influencing strength of parent-to-child nodes is shown in Table 2.6, where training has the highest weight for Operator Error with a value of 0.293. Based on the updated

BN model, the range of HEP under the organizational evidence factor can be gained, as Table 2.7 shows.

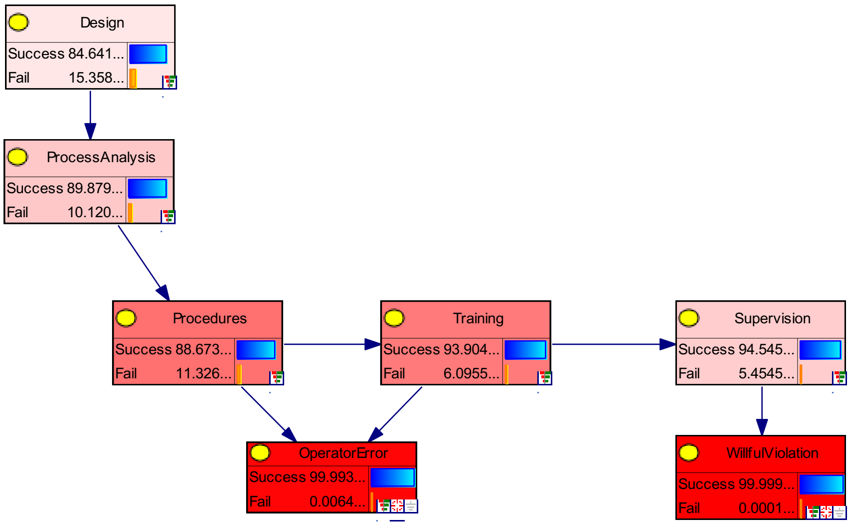


Fig. 2.16 The influencing strength of parent-to-child nodes

Table 2.6 The influencing strength of parent-to-child nodes

Parent	Child	Average	Maximum	Weight
Design	Process Analysis	0.36	0.36	0.36
Procedures	Training	0.391	0.391	0.391
Procedures	Operator Error	0.2095	0.262	0.2095
Process Analysis	Procedures	0.352	0.352	0.352
Supervision	Willful Violation	0.107	0.107	0.107
Training	Supervision	0.298	0.298	0.298
Training	Operator Error	0.2925	0.345	0.2925

Table 2.7 The influencing strength of parent-to-child nodes

Organizational factor	State	P(H2=Fail)	P(H4=Fail)	HEP
Procedure	0	3.9E-3	1E-4	4E-3
	1	2.61E-2	2E-4	2.63E-2
Training	0	4.5E-3	1E-4	4.6E-3
	1	3.67E-2	5E-4	3.72E-2
Supervision	0	5.9E-3	1E-4	6E-3
	1	1.58E-2	1.3E-3	1.71E-2
Process analysis	0	1.4E-2	1.3E-4	1.41E-2
	1	2.7E-2	1.3E-4	2.71E-2
Design	0	1.5E-2	1.3E-4	1.51E-2
	1	2.0E-2	1.3E-4	2.01E-2

2.5 Discussion

The analysis reveals significant differences in the "Who" and "When" dimensions of Human and Organizational Factors (HOFs)-related cases versus other cases. Tamim et al. [34] emphasize the pivotal role of engineering contractors in risk management across the lifecycle of a plant, drawing insights from the investigation of nine significant process safety incidents. Similarly, Zarei et al. [35] identify maintenance, control room, and emergency operations as critical for process safety. This study's findings support these perspectives, statistically demonstrating a higher incidence of HOFs-related accidents involving contractors that predominantly occur during maintenance phases. This contrasts with other scenarios, which exhibit lower occurrence rates. Consequently, this research will examine the potential of robotic implementations in maintenance tasks to mitigate these risks.

Regarding the "Why" aspect, the analysis indicates that over two-thirds of the identified factors are organizational elements. This provides evidence of the influence of organizational factors, which may serve as the root cause of both human and technical failures. One branch of this study goes deeper to develop a new Human Error Probability (HEP) calculation tool employing Bayesian Networks (BN). This BN is designed to establish a robust framework that can serve as the core model, with the flexibility to incorporate additional complexity as more data becomes available. The forthcoming chapter will delve into a comprehensive review of the current literature on risk assessment and Human Reliability Analysis (HRA) methods to anchor the proposed model within the state-of-the-art in this field.

In addressing the first research question (RQ1), the findings pinpoint procedures, maintenance, design, and training as the most critical organizational factors across all HOFs-influenced cases. Specifically, the paramount factors within the context of maintenance include organizational inadequacy, procedures, design and Human-Machine Interface (HMI), supervision and staff, and training. This delineation informs targeted interventions and policy formulations to reduce HOFs-related incidents, focusing on enhancing organizational practices and safety protocols.

Chapter 3

Literature Review

In one direction, substantial efforts have been made since 1980 to introduce more methods of human reliability analysis that can be summarized within a conceptual framework pattern. In another direction, the mainstream of risk assessment methodology research has expanded probabilistic risk assessment to a systemic approach; however, methods are still missing to represent the dynamic interdependence of the system elements. Therefore, the complex system theory and agent-based method and simulation(ABMS) will be employed in this research to better model HRT systems.

3.1 Human Reliability Analysis Methods Review

3.1.1 Key elements of HRA methodology

As a part of probabilistic risk analysis (PRA), human reliability analysis (HRA) is a tool to systematically identify and investigate the cause, consequences, and contributions of human failure in systems [43]. As a holistic methodology, the main components of an HRA should include the following:

1. A conceptual framework about human functional activities in the system, human error mechanism, and human error modes.
2. Implementation procedures and tools for identifying and analyzing the dependency of HFEs.

3. Basic data to quantify the human failure probability (HFP). Therefore, this research selects and reviews five holistic methodologies focusing on these three aspects.

So, this research will review six HRA methodologies according to these three aspects: Technique for Human Error Rate Prediction(THERP) [19], Cognitive Reliability and Error Analysis Method (CREAM) [20], Connectionism Assessment of Human Reliability (CHAR) [21], Information, Decision, and Action in Crew Context (IDAC) [22], A Model-based Human Reliability Analysis Methodology (PHOENIX) [44], and the General Methodology of An Integrated Human Event Analysis System (IDHEAS-G) [45].

3.1.2 THERP

Conceptual framework

1. Human activities include internal inputs (perception, discrimination), cognitive activities (interpretation, diagnosis, decision-making), and response (action).

2. Human error in the man-machine system is impacted by performance-shaping factors (PSFs), including external and internal factors.

3. Human error modes are simplified to only concerned with the outputs action part, Errors of omission (EOO), and Errors of commission (EOC).

Implementation procedures and tools

1. Define the system failures of interest that require personal actions related to system functions and for which error probabilities are estimated.

2. List and analyze the related human operations.

3. Utilizing an event tree method to estimate the relevant event structure and error probabilities, the effects of human errors on system failure events are assessed. This analysis considers event dependency and recovery factors to evaluate human errors' impact on the system comprehensively.

Primary data and source

1. 27 Data tables of nominal HEPs, PSFs, dependence, uncertainty bounds, and quantitative recovery factors. (Only 2 PSFs quantitative levels are given; others need experts' judgment.)

2. datasource: Nuclear power plants, Dynamic simulators of NPPs, Process industries, Job situations in other industries, and military situations are psychologically like NPP tasks.Experiments and field studies using real-world tasks of interest, e.g., experimental comparisons of measurement techniques in an industrial setting, performance records of industrial workers, etc. Experiments using artificial tasks, e.g., typical university psychology studies, have limited application to real-world tasks[19].

3.1.3 CREAM

Conceptual framework

1. Human cognition activities include Observation/identification, Interpretation, Planning/choice, and Action/execution

2. Human erroneous actions must include three main sets of factors or influences: a combination of individual, technological, and organizational factors. It is possible to define a small set of common performance conditions (CPC), which contains the general determinations of performance, including Adequacy of organization, Working conditions, Adequacy of MMI and operational support, Availability of procedures/plans, Number of simultaneous goals, Available time, Time of day (circadian rhythm), Adequacy of training and experience, Crew collaboration quality[20].

3. Systematic human error modes:

   (a) space dimension: direction, wrong direction, magnitude, too short, too far, object, wrong object

   (b) Time dimension: duration, too short, too long, timing, too early, too late, speed, too slow, too fast, sequence, the Wrong object observed

   (c) Force dimension: Force, too soft, too hard, unsteady

Implementation procedures and tools

Retrospective analysis:

1. Determine or describe the context.
2. Describe the possible error modes.
3. Describe the probable causes.
4. Perform a detailed analysis of the main task steps:
   (a) Describe initiating event
   (b) Identify error mode
   (c) Find the associated antecedents
   (d) Match antecedent with other classification groups

Performance prediction analysis:

1. Task analysis
2. Context description, using the common performance factors
3. Specification of the initiating events
4. Qualitative performance prediction
5. Selection of task steps for quantification
6. Quantitative performance prediction

Basic data and source

The level and dependency between CPCs The quantitative level of Control modes.

3.1.4 CHAR

Conceptual framework

1. Combined the definition of human error from the view of the philosophy of the concept of cognitive and procedure of information process, the view of psychology includes the cognitivism and behaviorism approaches, and the view of engineering sciences, human in a technical system, the new role of a human in the more and more automatic technical system. Give the definition" a human error always exists in a working system and is characterized by an undesired or faulty state of the working system. It then leads to a situation where the system's requirements are not met or are met inadequately. The individual is only one part of the working system and interacts together with other portions of the working system. All portions within the working system may depend on each other or maybe in a reciprocal action state"[21]. This definition emphasizes human error as defined by the requirements of the working system, and it is a combination of both load and cope aspects.

2. Categorize error models into phenomenological error models(error of omission, error of commission), causal error models(information processing error), and actional error models(in the form of PSFs).

3. Phenomenological Human error states:

   (a) Omission

   (b) Fault action commission

   (c) Human error

   (d) Human mistake

   (e) Error of confusion

   (f) Sequence error

4. Causal error description:

   (a) Task error

   (b) Execution error

   (c) Ergonomic error

   (d) System ergonomics error
   (e) Errors in certain situational conditions or environmental conditions
   (f) Perception(neglected)
   (g) Inadequate feedback(too little)
   (h) Disturbing environmental factors(too much)

5. Actional error description:
   (a) MMS component task
   (b) MMS component operator
   (c) MMS component activity and operation
   (d) MMS component feedback
   (e) MMS component technical system
   (f) MMS component order issue and order dispatch

Implementation procedures and tools:

1. Analysis of sources on practical operational experience
2. Analysis of data with a view to qualitative indications
3. Identified action types
4. Identified error types
5. Causal performance shaping factors
6. Identified error conditions or actional PSF
7. Interrelationships of error conditions
8. System ergonomics, cognitive errors, and errors of confusion
9. Frequency analysis-based predictions from practical operational experience
10. Assumption of distribution
11. Quantitative predictions based on practical operational experience

12. Estimated values for reliability parameters from practical operational experience with the help of a probabilistic model

13. Probabilistic predictions from practical operational experience

Primary data and source:

One hundred sixty-five events featuring human error data from the Society for Systems and Reactor Safety databank. All events from 1965 to 1993 and only events in boiling water reactors were considered. Give the value of errors of confusion, general, connected with maintenance tasks, when labeling is in effect, when clarity is in effect, when an arrangement is in effect, when time pressure is in effect, incorrect response to the occurrence of a latent error faulty quality control in connection with maintenance activities on several redundancies as error initiators[21].

3.1.5 IDAC

Conceptual framework:

1. Integrating the influences of rational and emotional dimensions involves formulating a concise set of general behavior rules that dictate the dynamic responses of the operator.Response:
   (a) Information pre-processing
   (b) Diagnosis and decision-making
   (c) Action execution
2. Mechanism:
   (a) Mental state
   (b) Memorized information
   (c) Physical factors
   (d) Intrinsic characteristics
   (e) Environmental factors
   (f) Conditioning event

   (g) Organizational factors
   (h) Team-related factors

3. Human error modes:
   (a) Perception stage: Miss transient indications, Delay in perceiving indications
   (b) Comprehension stage: Indication perceived but ignored, Incomplete use of information; investigation interrupted, Information decays from memory due to prolonged inactivity, Failure to retrieve relevant knowledge, Lack of knowledge[46]
   (c) Decision Stage: Jumping into a plausible false conclusion early, Waiting too long before decision-making, Misdiagnosing the situation due to plausible symptoms[46].
   (d) Error in action execution

3.1.6 PHOENIX

Conceptual framework:

1. Human response model:
   (a) Noticing/detecting/understanding
   (b) Situation assessment/diagnosis
   (c) Decision-making/response planning
   (d) Action taking
2. Types of crew activities:
   (a) Monitor
   (b) Scan
   (c) Detect/observe
   (d) Identify
   (e) Communicate

(f) Evaluate/interpret

(g) Record

(h) Compare

(i) Verify

(j) Adapt

(k) Adhere

(l) Diagnosis

(m) Decide

(n) Plan

(o) Coordinate

(p) Execute

(q) Regulate

(r) maintain

(s) Human error in the man-machine system is impacted by performance shaping factors (PSFs), including:

    i. Procedures
    ii. Resources
    iii. Team Effectiveness
    iv. Knowledge/Abilities
    v. Bias
    vi. Stress
    vii. Task load
    viii. Time constraint

(t) Human error modes are simplified to only concerned with the outputs action part, Errors of omission (EOO) and Errors of commission (EOC).

Implementation procedures and tools:

1. PRA scenarios development

2. Development of crew response tree

3. Identification of crew failure modes for CRT branches
4. Construction of HFE scenarios
5. Analysis of HFE scenarios, development of narratives, and identification of dependencies

3.1.7 IDHEAS-G

Conceptual framework:

Human activities include five macro-cognitive functions:

1. Detection cognitive activities
   (a) Detect cues
   (b) Acquire(gather) information
2. Understanding cognitive activities
   (a) Maintain situational awareness
   (b) Assess status based on indirect information
   (c) Diagnose problems and resolve conflicting information
   (d) Make predictions or form expectations for the upcoming situation development
3. Decision-making cognitive activities
   (a) Make a go/no-go decision for a pre-specified action
   (b) Select among multiple options or strategies
   (c) Change or add to a pre-existing plan or strategies
   (d) Develop a new strategy or plan
4. Action execution
   (a) Execution of a cognitively simple action
   (b) Execution of a cognitively complex action

   (c) Long-lasting action
   (d) Control action
   (e) Fine motor action
   (f) Physically strenuous action

5. Inter-team coordination
   (a) Communication
   (b) Cooperation
   (c) Coordination

Implementation procedures and tools:

1. Scenario analysis
   (a) Develop scenario narrative
   (b) Develop scenario context
   (c) Identify HFE
   (d) Define HFE
2. Modelling of critical human actions
   (a) Analyze tasks and identify CTs in HFE
   (b) Characterize the CTs and select applicable CFMs
   (c) Assess PIFs applicable to every CFM
3. HEP quantification
   (a) Calculate Pc
   (b) Analyze HFE timeline
   (c) Estimate parameters of Tavail and Treqd distributions
   (d) Calculate Pt
   (e) Calculate overall HEP
4. Integrative analysis, Uncertainty and dependency analysis, and documentation

Primary data and source:

The data available for this methodology is given in IDHEAS-EDA, which mainly comes from experience in the control room operation and training data. The data contains the basic value of the five macro functions based on the three most essential PIFs: task complexity, familiarity, and information reliability. Then, other PIFs will be evaluated to perform as multipliers to adjust the primary value for HEP.

3.2 Systemic Risk Assessment Methods Review

3.2.1 STAMP model

The STAMP (Systems-Theoretic Accident Model and Processes) model conceptualizes safety as a control issue, a paradigm shift introduced by Leveson [47]. Unlike traditional models that often attribute accidents to individual component failures, STAMP posits that the core cause of accidents is the failure to adequately control the system to ensure its operations remain within safe boundaries. Consequently, the creation of a robust control structure to regulate the system's operational processes is deemed crucial for accident prevention. This control structure may encompass several variables, and regardless of whether the control is executed by an automatic system or monitored by human operators, three critical pieces of information are essential: the current state of the system, the interrelations among system variables, and the strategies for modifying the system's state. Expanding the control structure diagram to include these elements—initial value, current state, and methods for state alteration—facilitates the development of a comprehensive control model.

Employing the STAMP model for accident analysis encompasses several key steps:

Identify Safety Requirements and Constraints: Initiate with a preliminary hazard analysis to discern potential hazards within the system that could lead to personal injury, equipment damage, or environmental harm. Following this, safety requirements and constraints aimed at mitigating these hazards must be established.

Define the Safety Control Structure: With the safety requirements and constraints identified, outline the safety control structure. This includes detailing the elements involved in safety constraints, control actions, and feedback mechanisms.

Identify Anomalies in Control Mechanisms: Examine the control and feedback paths to uncover potential anomalies or deficiencies that might lead to unsafe conditions or accidents.

Ascertain the Root Cause of Abnormal Control Actions: Through a thorough safety assessment of the system, analyze instances of abnormal control within the context of the identified safety constraints. This analysis should lead to the identification of root causes and the formulation of recommendations for system improvement.

The STAMP model offers a systems-theoretic perspective on accident prevention, emphasizing the importance of understanding and controlling the complex interactions within a system's operations. By focusing on these interactions rather than isolated failures, the model provides a framework for comprehensively addressing the multifaceted nature of system safety.

3.2.2 STPA model

STPA (Systems-Theoretic Process Analysis) is a methodology that scrutinizes the comprehensive risk controls and feedback loops within a system's safety control structure to unearth potential flaws or failures in these mechanisms that could precipitate accidents. Originating from Leveson's work [47], STPA is grounded in the understanding that safety issues are systemic and often arise from complex interactions within the system's control structure rather than isolated component failures.

The control structure model utilized in STPA outlines the system through a series of hierarchical levels that echo the framework proposed by Rasmussen, focusing on the intricate web of control and feedback relationships that interlink these levels. This hierarchical perspective facilitates a systematic examination of how decisions and actions at different levels of the organization or system influence overall safety.

To pinpoint potential risks stemming from inadequacies in the control mechanisms, STPA employs a taxonomy of control failures. This taxonomy serves as a tool to categorize and understand the diverse types of control flaws that can exist, ranging from issues in software and system design to human behavior anomalies, including those related to management and organizational factors. By identifying these potential control failures, STPA enables the mapping out of causal factors and hazardous scenarios that might not be immediately apparent through traditional risk assessment methods.

STPA thus offers a nuanced approach to safety analysis, emphasizing the identification and mitigation of systemic vulnerabilities within the control structure of complex systems. It extends beyond conventional hazard identification methods by considering how errors in decision-making, communication, and feedback can lead to unsafe conditions, thereby providing a more holistic view of system safety and risk management.

3.2.3 FRAM model

The Functional Resonance Analysis Method (FRAM) represents a paradigm shift in accident analysis, rooted in system theory and proposed by Hollnagel [48]. Its application spans diverse fields including aerospace [49] [50], construction [51], maritime [52], and nuclear energy [53], demonstrating its versatility and effectiveness in analyzing complex socio-technical systems.

FRAM's foundation is a systems approach to understanding accident causation and prevention, focusing on the functions of a socio-technical system rather than its structural components. It seeks to capture the dynamic nature of systems by illustrating nonlinear interactions and inherent performance variability [54]. Central to FRAM is the concept of functional resonance, which suggests that accidents occur not due to isolated failures but from the unexpected convergence of normal performance variability.

The method is built on four guiding principles [48]:

1. Failures and successes stem from the same origins.
2. The performance of socio-technical systems is adaptive aligned with current conditions.

3. System functions and their interrelations should be analyzed as they manifest in real-world scenarios, employing functional resonance to recognize how performance variability can aggregate unexpectedly.

The theoretical underpinning of FRAM, functional resonance theory, emphasizes the role of normal fluctuations within system functions and how abrupt deviations in these fluctuations can lead to accidents. Unlike traditional models that focus on how a function fails, FRAM concentrates on the conditions facilitating accident occurrence.

Implementing FRAM in accident analysis involves the following steps:

1. Defining System Functions: This initial step involves identifying the system's essential functions and representing them through a hexagonal function module, incorporating time, control, precondition, and resources alongside standard input and output parameters. This helps describe accidents beyond simple causality.

2. Characterizing Potential Variability: The FRAM model identifies eleven Common Performance Conditions (CPCs) that encompass the human, technological, and organizational aspects contributing to variability. These conditions range from personnel availability to organizational quality, setting the stage for understanding the bounds of performance variability.

3. Assessing Functional Resonance Possibilities: This involves examining how changes in functions might interact, establishing both expected and unexpected connections within a functional network. The focus is on identifying potential pathways through which these functions might influence one another under various conditions.

4. Implementing Protective and Control Barriers: Based on performance variability characteristics, FRAM suggests implementing protective and control barriers, which can be physical, functional, symbolic, or invisible. These barriers are designed to mitigate the risk of functional resonance by managing performance variability.

FRAM's holistic approach offers a nuanced perspective on system safety, emphasizing the importance of understanding and managing the complex interplay of functions and their variability. It provides a comprehensive framework for identifying potential pathways to accidents and devising strategies to enhance system resilience and safety.

3.2.4 "2-4" models

The "2-4" model proposed by Fu et al. (2013) [55] offers a nuanced framework for understanding accident causation, drawing upon extensive analyses of various accident causation chain models. This model delineates the evolution of accidents through four sequential stages reflecting both individual and organizational development: one-off behavior, habitual behavior, operational behavior, and guide behavior. Each stage represents a progression in the behaviors and practices that can culminate in accidents, emphasizing the dynamic nature of accident causation that spans from individual actions to organizational influences.

Central to the "2-4" model is the assertion that the immediate causes of accidents can be traced back to two primary sources: unsafe human behaviors and unsafe equipment and environmental conditions. This duality acknowledges the interplay between human factors and the physical context in which they operate, highlighting the complexity of securing safety across different scenarios.

The model further identifies the root cause of accidents as a deficiency in the organizational safety culture. This perspective shifts the focus from the symptoms of safety issues (unsafe behaviors and conditions) to the underlying systemic problems that enable such conditions to persist. By addressing the core aspects of organizational culture that contribute to safety lapses, the model suggests a more holistic approach to accident prevention.

Additionally, the "2-4" model introduces the concepts of internal and external influence chains, providing a comprehensive view of the accident causation process. The internal influence chain encompasses accidents arising from unsafe actions or conditions within the organization, driven by the behavior of its operators. In contrast, the external influence chain captures the sequence of accidents initiated by factors outside the organization, such as regulatory bodies, supervisory entities, and advisory organizations. This distinction between internal and external influences underscores the multifaceted nature of accident causation, recognizing the role of both internal practices and external pressures in shaping safety outcomes.

By integrating these elements, the "2-4" model presents a layered and systemic approach to understanding and addressing accident causation. It underscores the importance of fostering a robust organizational safety culture and acknowledges the complex interdependencies between individual behaviors, organizational practices, and external influences in the genesis of accidents.

3.2.5 Accimap

The Accimap approach represents a significant paradigm shift in the analysis of socio-technical systems, moving away from traditional methods that decompose these systems into discrete levels for multidisciplinary research. Instead, it adopts a cross-disciplinary, system-oriented perspective that focuses on examining the interactions among actors across all levels involved in managing hazard sources within specific workplace contexts. This approach underscores the importance of understanding the complex relationships and influences contributing to system safety and accident causation.

Unlike traditional models that dissect systems into their component parts (structural decomposition), the Accimap method emphasizes functional abstraction. This means that the system is conceptualized in terms of its functions and the interactions between those functions rather than its physical or organizational structure. The goal is to develop a system modeling language capable of capturing and representing performance and interactions across all system levels simultaneously. This approach allows for a more holistic understanding of how various elements within a socio-technical system interact and influence each other, particularly in the context of hazard management.

The Accimap method is particularly valuable in identifying and analyzing the multifaceted and often indirect paths through which systemic factors contribute to accidents. By mapping out the relationships and feedback loops between different levels of the system (e.g., government policy, organizational processes, workplace conditions, and individual actions), analysts can identify where failures or deficiencies may occur and how they propagate through the system. This comprehensive view facilitates the identification of leverage points for intervention and the development of more effective strategies for preventing accidents and enhancing overall system safety.

In essence, the Accimap approach provides a powerful tool for dissecting the complexity of socio-technical systems, offering insights into the dynamic interactions that underlie safety-critical processes. Focusing on functional abstraction and the system as a whole enables a deeper understanding of the systemic roots of accidents and the broader context in which they occur.

3.2.6 IDDA

The IDDA methodology is predicated upon an enhanced iteration of an event tree [56], underpinned by a logical-probabilistic modeling framework that fuses the system with its phenomenological model [57] [58]. This approach could represent event inter-dependencies via logical constraints, thereby imparting the capacity to modify the structural configuration of the event tree, influence its outcomes at varying hierarchical levels, or revise the real-time probabilities associated with event occurrences. The IDDA method has been extensively validated in numerous instances of process risk assessment [59] [9]. Consequently, in the context of this research, IDDA is employed.

3.3 Complex System Theory and Methods Review

3.3.1 Complex system theory

The conception framework of complex systems was proposed by a group of scientists at the Santa Fe Institute [60]. This theory originated from the field of biology and then provided guidelines about the common phenomena seen in many different areas of research, such as economics [61], political sciences [62], ecosystems [63], education [64], medical science [65].

Complex systems consist of heterogeneous functionally integrated interacting subsystems and agents in a nonlinear fashion that can adapt and learn, enabling these agents to achieve collective properties that individual one does not have [66] [67] [68] [69] [70]. In nonlinear relationships, changes in the input to the system do not lead to direct proportion to changes in the output [71]. Therefore, the complex system theory introduced a holistic paradigm other than the traditional reduction paradigm when selecting research approaches that are better suited to represent the HRT systems.

3.3.2 System dynamic simulation

The system dynamic(SD) method employed a coupled series of ordinary differential equations to represent the system. The main steps include mapping out causal loops, identifying state variables of interest, identifying flows of interest, and defining supporting variables [72]. SD has been proven to be a powerful tool for analysis of the underlying dynamics of the systems, offering valuable support for decision-making. SD Start with the entire system. Assumes that the analyst knows how the system behaves and encodes their understanding.

3.3.3 Agent-based modeling and simulation

Agent-Based Modeling and Simulation (ABMS) represent a bottom-up methodology designed to capture emergent phenomena occurring at the systemic level as a result of interactions among individual entities. ABMS is acknowledged for its ability to accurately capture individual behaviors, particularly when they exhibit nonlinearity, memory path-dependence, and non-Markovian characteristics [73].ABMS scholars propose that the initial macro conditions shape and drive individual actors' actions. These interactions subsequently accumulate to form a new macrosocial outcome. Through a bottom-up approach, the focus lies on human actions to uncover the underlying mechanisms linking social factors [74].

In a study by Abdelkhalek and Zayed [75], an agent-based and discrete event simulation hybrid model was developed to investigate the temporal and cost aspects of the inspection process for concrete bridge decks employing non-destructive technologies. While this research considered various process plan details, encompassing the selection of different machinery and inspection technologies, it did not encompass an evaluation of inspection result accuracy or safety operation aspects. Janssen et al. [76] utilized an agent-based modeling and simulation approach to conduct security risk analysis within airport operations, taking into account temporal and spatial factors. However, this study did not incorporate considerations of human-machine interactions or machine failures.

In a separate study by Guo et al. [77], a reliability modeling framework featuring a continuous description of system performance, inclusive of a management agent and various system agents with diverse normal and failure-type states, was developed. This framework was applied in a case study involving subsea oil-extracted infrastructure. Notably, the study did not delve into the intricate realm of human factors. Wu et al. [78] developed an agent-based multi-fidelity modeling approach for the Human-Robot Collaboration (HRC) process, primarily investigating the impact of human-robot collaboration on construction productivity. While their study introduced three critical human factors—forgetting, muscle fatigue, and behavior uncertainty[78]—it did not account for pivotal factors such as robot failure and routine maintenance interruptions. The omission of these factors may potentially lead to an overestimation of the efficiency of human-robot collaboration scenarios.

The aforementioned studies collectively underscore the advantages of applying ABMs theory to complex system analysis. In the context of this research, ABMS make distinctive contributions in three significant domains. Firstly, they expand the scope from regular operational scenarios to encompass failure scenarios, thereby considering human error, robot failures, and maintenance processes. Secondly, these models incorporate dynamic human failure reliability analysis methodologies. Lastly, they facilitate a comprehensive comparison between traditional full manual inspection systems and Human-Robot Team (HRT) inspection systems, focusing on efficiency, accuracy, and safety aspects.

According to ABMS, distinct boundaries separate individuals from one another. These individuals exhibit behaviors and engage in actions over time. Observables represent measurable characteristics of interest, which can be linked to individual entities or the entire group. Typically, the values of these observables evolve over time. Agent-based modelers primarily focus on modeling behaviors, while system dynamics (SD) modelers typically begin with variables and equations. In this research, human behavior needs to be represented and measured. Therefore, the ABMS, other than SD, is selected to model the complexity of the interdependency of the system elements.

3.4 Discussion

As analyzed in Section 3.1.3, the five organizational factors (discussed in Chapter 2) could be expressed as CPCs through the CREAM method, and this method also contains the basic elements of the HRA conceptual framework. Therefore, the HRA method selected for this research is the CREAM method with some extensions. This could conquer the first challenge for integrated framework building(discussed in Section 1.3).

After the literature review of the main tendency in the development of systemic PRA methods, the IDDA method shows its priority in its exhibition of all the possible consequences and dynamic logic constraints; in addition, the IDDA software tool is efficient at calculating the probability of events and changes. However, IDDA could not represent the human-robot communication mechanism dynamically. Therefore, under the umbrella of complex system theory, the ABMS is selected to give a more detailed representation of human behavior. In this way, the second challenge will also be solved, and the integrated framework will be demonstrated in the next Chapter.

Chapter 4

The Integrated Framework

Based on the accident data analysis (in Chapter 2) and literature review (in Chapter 3), an integrated framework is proposed, where the CREAM method is chosen to perform the HRA part with some extension. The IDDA and ABMS are chosen to represent the interdependencies of the system elements. Related work has been published in Safety Science, entitled"Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations" [26]. Another paper in Advanced Engineering Informatics entitled " Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis" is under review.

4.1 Design Principles

Based on the complex system theory and related work gaps analysis in Chapter 1 and Chapter 3, five principles were proposed to guide the integrated framework building.

4.1.1 Comprehensive

As mentioned in Chapter 1, the HRT system is a socio-technical complex system with multiple elements and interactions between them. Therefore, this framework must integrate failure risk original from multiple sources and represent the interactions between the elements, especially the risk original from human error.

4.1.2 Interaction

Unlike the traditional methods like fault tree. The agent-based model can represent the interactions between different elements by sending messages between agents to trigger the state change.

4.1.3 Dynamic

It is better to employ real-time system performance measurement to consider the interactions between the elements. The framework should include methods representing the changes in element and system state over time.

4.1.4 Uncertainty inclusive

Uncertainty originating from individual capacity differences will influence the system performance measurement. Therefore, in this framework, this kind of uncertainty should be represented by some probability distribution functions other than just the point value.

4.1.5 Combination of qualitative and quantitative methods

The framework should include both qualitative and quantitative methods. In this way, the results of the qualitative analysis could give better insight into the system failure and performance dynamic mechanism. In contrast, the results of the quantitative analysis could provide a better comparison in data analysis for the traditional full manual and HRT scenarios.

4.2 Overview Structure

Fig. 4.1 shows the proposed framework. Along with the traditional risk assessment techniques, the framework starts with a preliminary qualitative analysis phase aimed at analyzing system tasks, their interdependency, and related hazards. This qualitative phase consists of a top-down strategy from the system level to the elements level to capture the emergence at the system level. Then, a quantitative analysis phase is employed to quantify the system performance, including the probability of unwanted outcomes and their consequences. At the element level, the HRA method is selected, extended, and utilized to analyze human reliability (the selection of HRA methods will be discussed in 4.3)[26]. Data for robot reliability and equipment reliability are derived from the literature. At the system level, to get predicted risk-based system performance, the Integrated Dynamic Decision Analysis(IDDA) [56] method integrates the elements' reliability in a dynamic event tree and the logical interdependencies, identifying and assessing the potential outcomes. To model the agent behavior in the environment in a more detailed manner, the agent-based modeling and simulation(ABMS) method is utilized to simulate the system's performance. These two system-level methods could serve as cross-validation and complement each other.

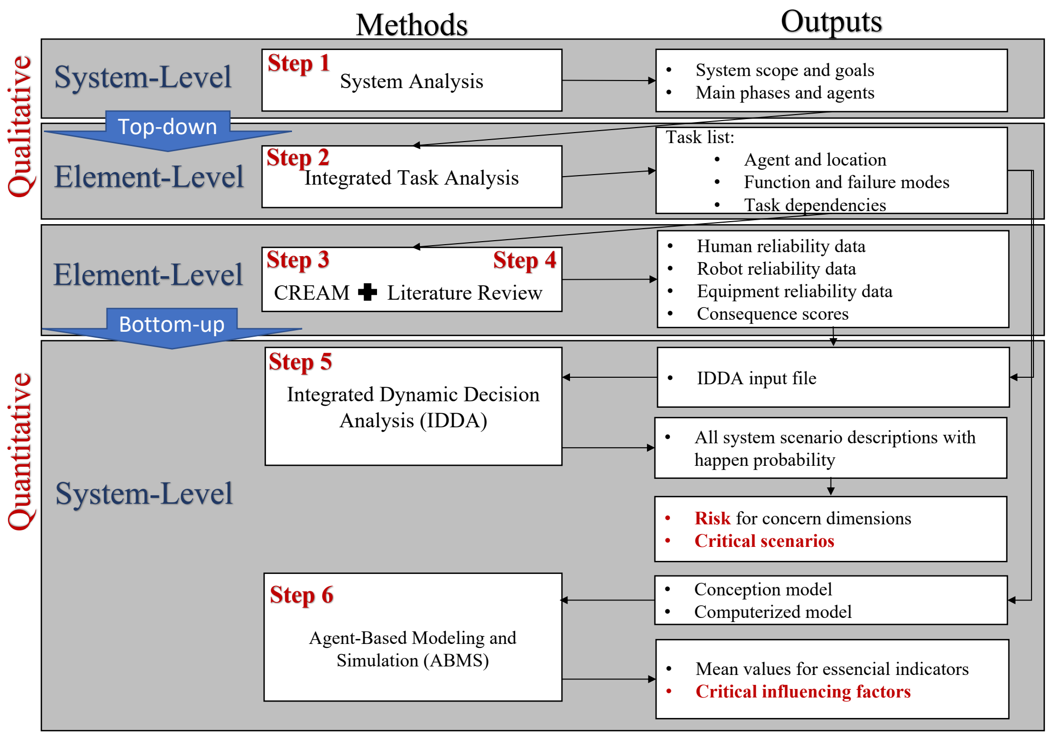


Fig. 4.1 Structure of the integrated framework

As shown in Fig. 4.1, the framework implementation includes 6 steps: system analysis, integrated task analysis, human reliability estimation, robot reliability estimation, IDDA, and ABMS.

4.3 Step1: System Analysis

The initial step in implementing the proposed framework involves identifying the system goals based on the stakeholders' expectations. Following this, it is essential to establish the framework's foundation by determining the boundary between the external environment, the target system, and the key elements[26]. Subsequently, the critical attributes, actions, and interactions of these elements should be specified.

4.4 Step2: The Integrated Task Analysis

Task analysis involves examining the actions or cognitive processes required by an operator or team of operators to accomplish a system goal [79]. In this research, task analysis is integrated with the cognitive functions for each task's demands[26]. Four functions are considered: Observation(O), Interpretation(I), Planning(P), and Execution(E) according to its matrix with cognitive activities [20], e.g., if a subtask activity is monitoring, then the demand cognitive function will be O and I, and the corresponding cognitive failure modes as miss an alarm while monitoring will be O3, as wrongly understanding an alarm will be I2. On the one hand, this structure is consistent with the state-of-the-art cognitive process models used in HRA methods, such as IDAC [22] and IDHEAS [45]. On the other hand, this structure can be extended to represent the robot functions as the HRI model built based on communication theory [80]. Task analysis can also identify hazards and consequences. This study incorporates these characteristics into the integrated task analysis to minimize procedural overlaps. An example of the main content of the integrated task analysis is shown in Table 4.1, including the column of actor, cognitive function failure mode(CFM), hazard, input, output, and consequence, which also can include location and tools columns if they have an impact on task performance[26].

Table 4.1 The integrated task analysis example

Task		Subtask	Actor	CFM	Hazards	Input	Output	Consequence
1 Set the robot	1.1	Connect the robot with a power strip	operator	E3	Electricity shock		Robot with power	Occupational Accident
	1.2	Set the robot parameters	operator	O,I,E	Forget the right value		Robot got set	Time delay
	1.3	The get-ready indicating light turns on	Robot	E		1.2	The signal sent by the robot	Time delay
	1.4	Observe the robot indicator	operator	O	Distraction	1.3	The signal received by the operator	Time delay

4.5 Step3: Human Reliability Estimate

Three things need to be considered when selecting an HRA method in the framework. Firstly, unlike the first-generation HRA methods that treat humans with input and output the same as a machine, the development of HRA methods considers human cognitive functions and the mechanisms that influence cognitive function failure. This improvement of the HRA could provide better explanations of human failure events and better serve the decision-makers. Secondly, the results of HOFs contributed to accident analysis in Section 2.5, showing that the contractors are an important influencing factor for this kind of accident. Also, in the organizational influencing factors analysis, the training and procedure factors show their priority. This emphasizes the significance of including inter-team communication, coordination, and training and procedure factors in the HRA method. Thirdly, one of the goals of this research is to explore the difference between the HRT and traditional fully manual system performance. Therefore, quantitative data is needed to make the comparison.

4.5.1 CREAM

As reviewed in 3, some HRA methods could already be used to calculate the HEP for a single event or function. Among them, the CREAM method could express the influence of organization and HMI in CPCs and consider the cognitive function modes of the event [20].

CREAM, as a representative second-generation method, has been widely applied and validated in the energy and chemical industry [81] [82], the maritime industry [83], and the transportation industry [84]. CREAM enables the quantification of human error probability through a two-step process. First, the initial error probabilities of the subtasks can be obtained based on the cognitive function failure modes derived from the integrated task analysis. For example, the initial failure rate for missing an alarm is categorized as O3, with a probability of 7×10^{-2}. Secondly, the multiplier selection is based on the level of Common Performance Conditions (CPCs), the assessment of which is performed through expert opinion aggregation[26].

4.5.2 DST-CREAM

One shortcoming of CREAM is not providing any guidance for aggregating assessment across experts [85]. To fill this gap, the Dempster–Shafer theory (DST) [86] [87], also called a “theory of evidence," is employed to fuse the expert’s opinions on the CPCs. As a generalization of both probability and possibility theories, DST has the capacity to reduce uncertainty in evidence from different sources in an effective and valid manner [88]. It is widely used in the process of combining evidence [89] [90].

Suppose the set consisting of all possible observed situations for variable X is called a frame of discernment (FOD) by $\Theta = \{H_1, H_2, \cdots, H_n\}$. The set contains finite and mutually exclusive hypotheses. The strength of it to a power set is represented as $2^{\Theta} = \{\varnothing, \{H_1\}, \{H_2\}, \cdots, \{H_n\}, \{H_1 \cup H_2\}, \cdots, \{H_1 \cup H_2 \cup \cdots H_i\}, \cdots, \{H_1 \cup H_2 \cup \cdots H_n\}\}$.

Each subset in the power set represents a proposition regarding X. A piece of evidence may support multiple propositions and a value between 0 and 1 can be assigned based on the degree of belief in the evidence. This is expressed through the basic probability assignment (BPA), also referred to as the mass function m: $2^{\Theta} \to [0, 1]$, and satisfies the Equation 4.1:

$$\begin{cases} m(\varnothing) = 0 \\ \sum_{X \in 2^{\Theta}} m(X) = 1 \end{cases} \tag{4.1}$$

Where X denotes one of the propositions in 2^{Θ} and is called the focal element if m(X)>0, suppose there are two mass functions, m_1 and m_2, from the same FOD. Fusing two BPAs collected from different sources can follow Dempster's combination rule, denoted by $m = m_1 \oplus m_2$. this is defined as Equation 4.2 and 4.3[91]:

$$m(X) = \begin{cases} \sum_{B \cap C = X} \frac{m_1(B)m_2(C)}{1-k}, X \neq 0 \\ 0, X = \varnothing \end{cases} \tag{4.2}$$

$$k = \sum_{B \cap C = \varnothing} m_1(B)m_2(C) \tag{4.3}$$

Based on the combined weight, the appropriate level of each CPC could be determined. Consequently, the corresponding multiplier for each cognitive function could be assigned accordingly. After that, the adjusted value for basic cognitive function failure modes will be calculated[26].

4.5.3 DDST-CREAM

To model human fatigue and recovery changing over time and its mathematics metric with human error probabilities. The fatigue-recovery human reliability model built by Givi et al. [92] will be employed to extend the DST-CREAM method to a dynamic one as DDST-CREAM. The formula for fatigue value calculation is as Equation 4.4,and the formula for recovery value calculation is as Equation4.5[92] :

$$F(t_i) = R(t_{i-1}) + (1 - R(t_{i-1}))(1 - exp(-\lambda \times t_i) \tag{4.4}$$

Where $F(t_i)$ means the accumulated fatigue value till the end of the i^{th} time cycle, $R(t_{i-1})$ means the residual fatigue value till the end of the $(i-1)^{th}$ time cycle, t_i means the time duration of the i^{th} time cycle, and t_{i-1} is the time duration of the one before the i^{th} time cycle, λ means the fatigue value index which describing the severity of the work performed.

$$R(t_i) = F(t_{i-1})exp(-\mu \times t_i) \tag{4.5}$$

Where $R(t_i)$ means the accumulated recovery value till the end of the i^{th} time cycle, $F(t_{i-1})$ means the accumulated fatigue value till the end of the $(i-1)^{th}$ time cycle, t_i means the time duration of the i^{th} time cycle, and t_{i-1} is the time duration of the one before the i^{th} time cycle, μ means the recovery value parameter.

The fatigue value starts from 0 at the work start point, and the fatigue value is supposed to be 1 after 4, 6, 8, or 12 hours of work. The fatigue index λ and recovery index μ can be calculated based on Equation 4.4 and 4.5, using the time unit as minute. Then the fatigue index λ will be 0.0288, 0.0192, 0.0144, or 0.0096, the recovery index values are likewise.

In the literature, the human error ranges from 1×10^{-6} to 1, then can be mapped with the fatigue value with the Equation 4.6:

$$Log(HEP) = 6Log(Ft_i) \tag{4.6}$$

Combine the Equation 4.7 and the CREAM method, the human error probability range could be obtained from the CREAM control mode, then the HEP can be mapped with the HEP_{base}, and the fatigue value with the Equation 4.7:

$$HEP = 10^{6*Log(w_f*F(t_i)+w_b*HEP_{base})} \tag{4.7}$$

where w_f and w_b are the weight for fatigue value and HEP_{base} , they are set as equal as an initial attempt.

To test the availability of this method, the plots of different values of fatigue index and HEP_{base} are shown in Fig. 4.2 and 4.3, with the point value at 120 min, which is the usual work shift duration. The HEP values vary in a rational range.

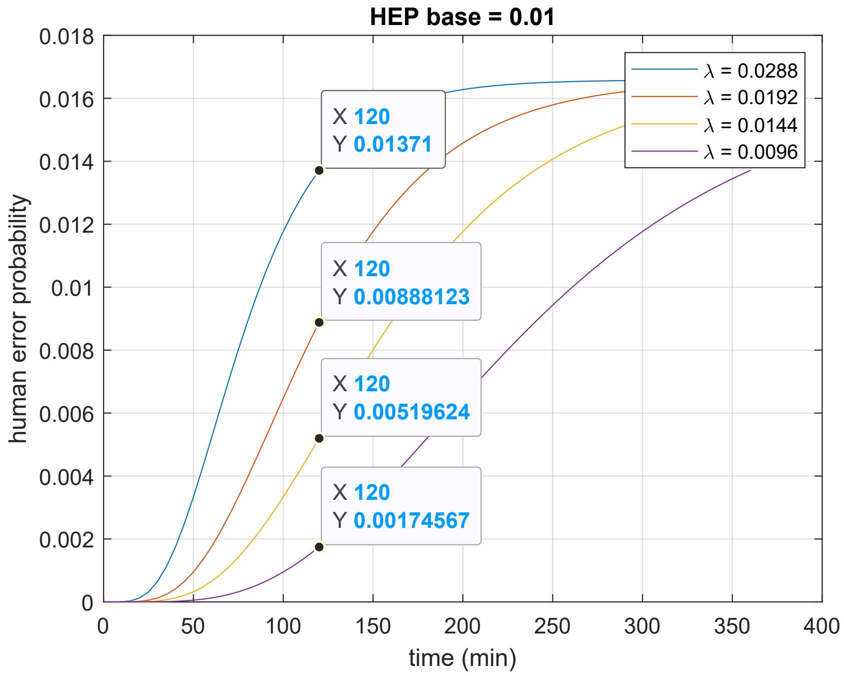


Fig. 4.2 The HEP range responding to the fatigue index values when $HEP_{base} = 0.01$

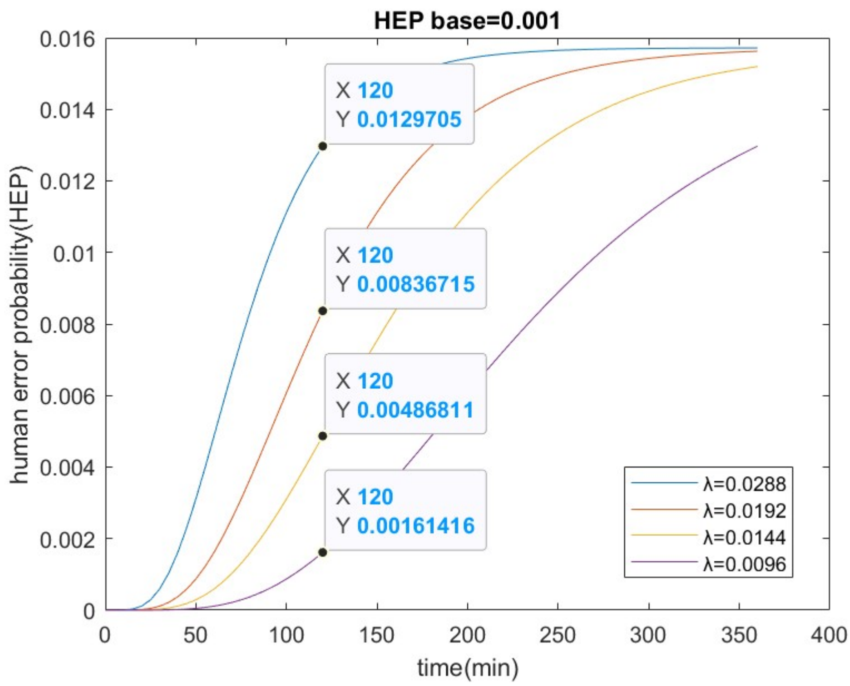


Fig. 4.3 The HEP range responding to the fatigue index values when $HEP_{base} = 0.001$

4.6 Step4: Robot reliability estimate

The robot is composed of several integral components, including a mechanical section, a control system, sensors, and actuators. The mechanical aspect, featuring a crawler mechanism for mobility on spherical surfaces, facilitates the robot's movement. Notably, the manipulation components, referred to as end actuators, play a crucial role in executing tasks such as polishing or magnetization. Actuators, encompassing servomotors, drivers, and transmission systems, provide power to various parts of the robot. Sensors, including cameras and radar sensors, serve to gather data from the mechanical system and provide environmental information[93].

In accordance with China's industry regulations pertaining to robots, the Mean Time Between Failures (MTBF) is stipulated to be more than 50,000 hours, leading to an estimated overall robot failure rate of 2E-6. Empirical research findings indicate that the distribution of failure rates among different components is as follows: the control system accounts for 32%, the end-effectors for 27%, communication for 16%, sensing for 12%, and power for 12% [94].

Based on the function mechanism of the robot, the sensor failure and the power failure contribute to observation function failure, the control system and the power failure contribute to the interpretation function failure, the control system the power failure contribute to the plan function failure, and the actuator failure, power failure and communication failure contribute to the execution failure[26]. Following the plus rule of the combination of independent parts failure rate, the basic robot function failure rates are estimated as Table 4.2 shows. A communication function is added to express the human-robot communication channel failure.

Table 4.2 Estimated failure rate for robot functions

Robot Function	Failure Rate Estimated
Observation (Power + Sensing)	4.8E-07
Interpretation (Power +Control system)	8.8E-07
Plan (Power + Control system)	8.8E-07
Execution (Power + End-effector +Communication)	1.1E-06
Communication	3.2E-07

4.7 Step5: IDDA

Upon the completion of individual probability assessments for each constituent element within the system, the consolidation of these assessments into a comprehensive overview is facilitated through Integrated Dynamic Decision Analysis (IDDA). The IDDA methodology is predicated upon an enhanced iteration of an event tree [56], underpinned by a logical-probabilistic modeling framework that fuses the system with its phenomenological model [57] [58]. This approach accommodates the incorporation of event inter-dependencies via logical constraints, thereby imparting the capacity to modify the structural configuration of the event tree, influence its outcomes at varying hierarchical levels, or revise the real-time probabilities associated with event occurrences. The IDDA method has been extensively validated in numerous instances of process risk assessment [59] [9]. Consequently, in the context of this research, IDDA is harnessed to amalgamate the developmental trajectories of diverse system scenarios. An input file can be created using the outputs from integrated task analysis, human reliability estimation, and robot reliability estimation. This file will consist of mutually exclusive sequences, each assigned specific occurrence probabilities. The IDDA method is fully equipped to comprehensively delineate the system's condition by all conceivable propagating paths generated from a system-specific input file. The input file employs logical constraints to express event inter-dependencies, as exemplified by instances such as: "L 101 1, 102 1," signifying that the failure of event 101 entails the concomitant failure of event 102; "A 101 1, 102 103 104," implying that the failure of event 101 results in the transformation of the success and failure state trajectories of event 102 into event 103 and event 104; and "P 101 1, 102 0.01," denoting that the failure of event 101 leads to a revision of the failure probability of event 102 to 0.01. A straightforward example of the IDDA modeling approach for specifying a risk model is illustrated in Fig.4.4. The IDDA method can fully represent the system state according to all the possible propagating paths generated from an input file developed on a specific system[26].

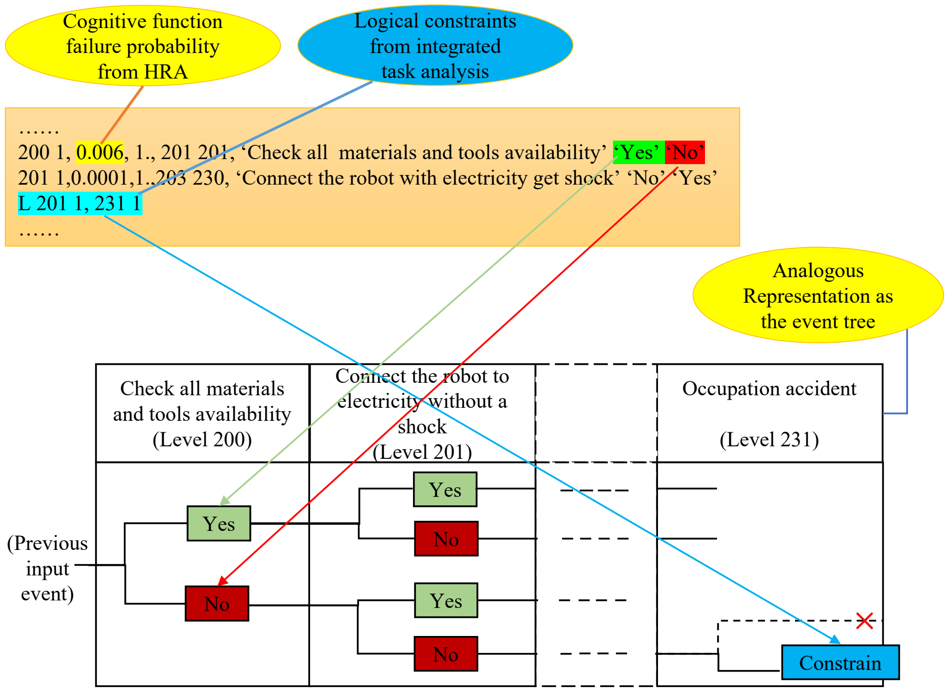


Fig. 4.4 Demonstrate the IDDA method

4.8 Step6: ABMS

The agent-based simulation method could model the human operator and robot as independent agents with attributes expressed by different parameters. This method is chosen to model the total time duration and cost performance of the HRT system, for it has three key advantages:

1. Using the real-time generated data from probability distribution functions other than the fixed-point value. In this way, the difference between individuals can be represented, and then the uncertainty that comes from the different capacities of the individual, such as the moving velocity or operation velocity, can be simulated.

2. Using the message-sending and receiving functions to simulate the process of human-robot or human-human communication and interaction.

3. Building the dynamic agent-based model simulates the agent attributes changing over time, such as the fatigue value and human failure rate.

4.9 Discussion

The proposed comprehensive, integrated framework includes qualitative and quantitative methods to explore the interactions between the organization, human, robot, and technical components and risk generation and propagation mechanism. This framework could compare the differences between the traditional full manual and HRT in terms of team structure and communication mode changes for the benefit of qualitative analysis and system performance changes in numeric form for the benefit of quantitative analysis; this will be tested in the case study in the next Chapter.

Chapter 5

Case Study

Fatigue cracking is one of the main contributors to the failures of engineering structures, which can lead to integrity loss [95]. Fluorescent Magnetic Particle Inspection (MPI) is an early and widely used inspection method for Non-Destructive Testing (NDT) of surface and near-surface flaws of metal parts [96] [97] [98]. It uses the leakage magnetic field formed at the surface defect to adsorb the magnetic suspension particles with fluorescent dyes so that they can accumulate at the crack. fluorescent MPI has a simple process, high inspection, high sensitivity, and is not affected by the size and shape of parts [99].

The cases introduced come from the Special Equipment Inspection Institute in Zhejiang, China. A periodic non-destructive detection involving defects test of weld joints of a pressurized spherical storage tank was selected as a case study to apply the methodology. The traditional process needs human operators to work a long time inside the spherical storage tank, the confined space of a pressured vessel, which causes risks to the process, occupational health, and work delay [100]. Therefore, comparing traditional and human-robot collaborative processes may show different risks. The procedure for MPI non-destructive testing of the spherical vessels consists of many steps, all of which must be captured in the task analysis. Related work has been published in Safety Science, entitled"Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations" [26]. Another paper in Advanced Engineering Informatics entitled " Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis" is under review.

5.1 Case description

The specific pressure vessel considered in this case study is a 5000 m^3 spherical, above-ground LPG storage tank. The total length of all weld joints is about 710 m. There are two polar zones of weld joints and two temperate zones of weld joints.

5.2 Data collection

The cases introduced originate from the Special Equipment Inspection Institute in Zhejiang, China. A periodic non-destructive detection test for defects in the weld joints of a pressurized spherical storage tank was selected as a case study to apply the methodology. Traditionally, this process requires human operators to work for extended periods inside the confined space of the spherical storage tank, posing risks to the process, occupational health, and potential work delays. Therefore, comparing traditional processes with human-robot collaborative processes may reveal different risk profiles. The procedure for fluorescent magnetic particle inspection (MPI) non-destructive testing of the spherical vessels involves numerous steps, all of which must be included in the task analysis. The specific pressure vessel in this case study is a 5000 cubic meter, above-ground spherical storage tank[26].

Three approaches were employed to ensure the reliability of the study:

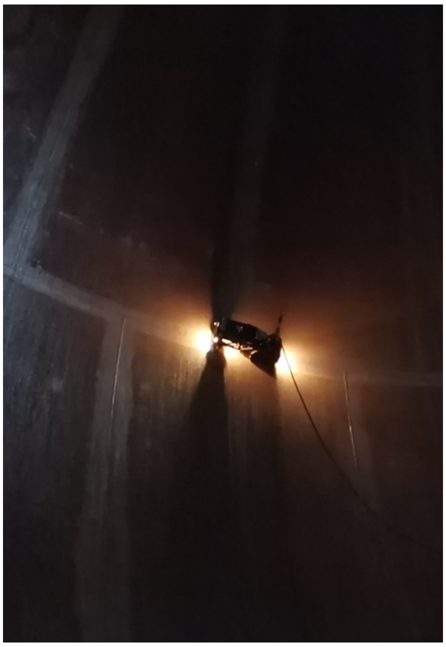
1. Semi-structured interview. Information necessary to describe the fluorescent MPI process on site was gathered through semi-structured interviews. Four field experts participated in these interviews: one project manager, one inspection team technician, and two robot testing technicians, all with 7-10 years of practical experience in fluorescent MPI projects. The interview script was designed to collect the following information:

    (a) The phases of fluorescent MPI for pressurized spherical tanks and their corresponding goals and subgoals.

    (b) The teams involved in each phase, their respective goals and functions, and the basic information to support CPCs assessment B.1.

    (c) critical task scenarios such as occupation risk and potential for incorrect results.

(a) FM field scenario



(b) HRT field scenario

Fig. 5.1 Field scenarios of the FM and HRT

2. Technical document collection. Relevant inspection documents were collected to provide details for implementing the method. These documents included the inspection organization plan, scheme, technical standards and guidelines, risk and hazard list, and emergency response plan.

3. Field investigation. A team of three scholars conducted a field investigation of a fluorescent MPI spherical tank inspection process. They observed and recorded both the fully manual scenario (shown in Fig. 5.1a)and human-robot (shown in Fig. 5.1b) scenarios, which were visited and recorded.

5.3 System analysis

Through the interviews with stakeholders, the main objective of the spherical tank inspection activity is to detect early faults correctly while considering occupational safety and time duration[26]. Therefore, system goals could be summarized into efficiency, accuracy, and safety, and the corresponding risk consequences, including

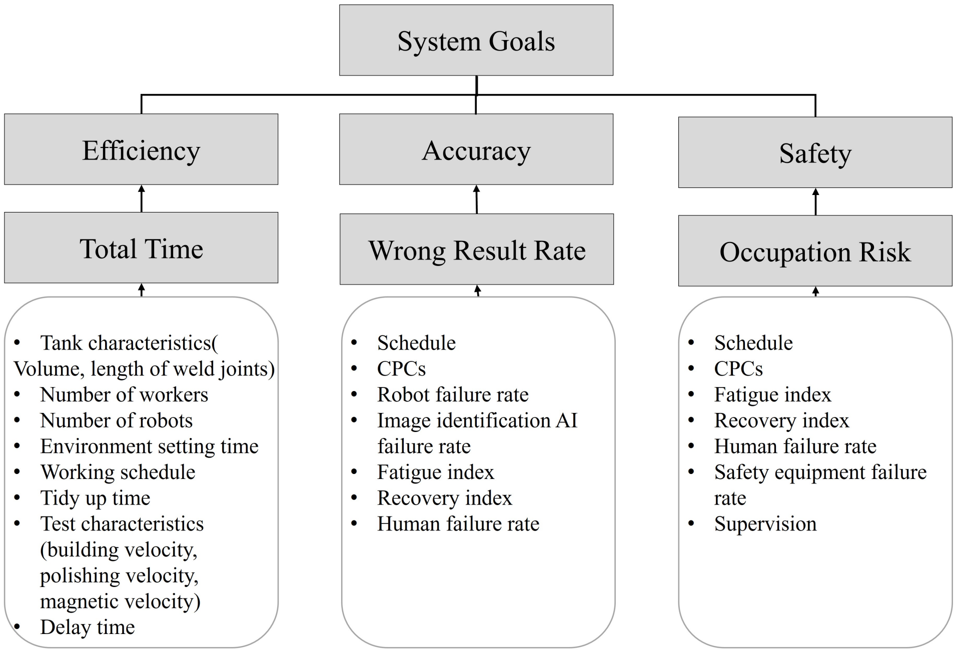


Fig. 5.2 Parameters influence system goals

time delay, wrong results, and occupational incidents, were analyzed in the integrated task analysis.

5.3.1 System goals

The main object of the spherical tank inspection activity is to detect the early fault correctly while considering occupational safety and time duration[26]. Therefore, system goals could be summarized into efficiency, accuracy, and safety. The parameters influencing system goals are shown in Fig. 5.2

For efficiency, the project's total time is chosen as the indicator, influenced by manual working time, human-robot collaboration time, robot working time, moving time, maintenance, and delay time. The formula for mean total time can be:

$$TT = T_{end,i} - T_{start,i} \tag{5.1}$$

Where TT represents the mean total time,$T_{end,i}$ represents the project end timestamp, and $T_{start,i}$ represents the project start timestamp.

The project's wrong result rate is chosen as the indicator for accuracy, influenced by schedule, human error, robot failure rate, and AI failure rate. The formula for the mean miss rate can be:

$$WR = (N_{miss} + N_{falsetrue})/N_{crack} \tag{5.2}$$

Where WR means the mean wrong result rate,N_{miss} is the number of missed cracks in the i^{th} experiment, $N_{falsepositive}$ is the number of false true cracks in the i^{th} experiment, and N_{crack} is the number of real cracks.

For safety, whether an occupational incident happens or not is chosen as the indicator and is influenced by schedule and human error. The formula for the occupational incident can be:

$$O_h = max(O_1, O_2, \cdots, O_n) \tag{5.3}$$

Where O_h means the occupational incident happens as 1 or not as 0, and O_i implies the type n incident happened incident as 1 or not as 0, n refers to the number of occupational incident type.

5.3.2 System phases

The FM inspection process includes six stages: environment setting(prepare), scaffold building, polishing, magnetic, scaffold rebuilding, and tidy-up. The HRT inspection process consists of four phases: environment setting, polishing, magnetic, and tidy-up. The environment setting and tidy-up phases are simplified in this research, for they only differ in duration for the HRT and the FM system. The scaffold-building and dismantling process will be simulated by working speed and workload in terms of the volume of the spherical tank.

The focus is on polishing and magnetic processes. To represent these two phases, human workers or robots individually move along the weld joints with the working speed. In addition, the robot's movement path and difference in velocity of moving and working will also be considered.

System functions failure and transition paths

The inspection process includes five main tasks and an emergency operation. The detailed analysis of 50 subtasks is shown in Table B.1:

1. Preparation and environment setting
2. Scaffold building
3. Manual polishing
4. Manual magnetic and crack identification
5. Scaffold dismantle
6. Emergency management

They are reduced to three and an emergency operation using robots. The detailed analysis of 45 subtasks is shown in Table B.2:

1. Preparation and environment setting
2. HRT polishing
3. HRT magnetic and crack identification
4. HRT emergency management

5.4 Human Reliability Analysis

Based on the integrated task analysis, the functions of human activities and their failure modes were identified. Using Table B.2, the nominal probability value of the cognitive failure mode was estimated. Subsequently, the Level of CPCs was evaluated. In this case study, five teams needed to set CPCs: the plant team (Team1), the construction team (Team2), the polishing team (Team3), the full manual testing technician team (Team4), and the human-robot testing team (Team5). Five experts in chemical engineering and human factors were invited to assess the weight of CPC levels based on the basic information of the five teams.

They were given the basic information about CPCs for organizations (shown in Table B.1). Utilizing Equations 4.2 and 4.3, the evidence from five experts can be fused, and the combined weights for each CPC level are shown in Table 5.1. After selecting the appropriate multiplier and combining it with the nominal value, the failure rate of each cognitive function failure mode for each team could be calculated. The results of these calculations are presented in Table 5.2.

Table 5.1 The fused weight value for levels of each CPC

CPCs	CPC1				CPC2			CPC3				CPC4			
Levels	L1	L2	L3	L4	L1	L2	L3	L1	L2	L3	L4	L1	L2	L3	
Team1	0.70	0.30	0.00	0.00	0.00	1.00	0.00	0.35	0.65	0.00	0.00	1.00	0.00	0.00	
Team2	0.26	0.74	0.00	0.00	0.00	0.91	0.09	0.35	0.65	0.00	0.00	0.00	1.00	0.00	
Team3	0.22	0.78	0.00	0.00	0.00	0.81	0.19	0.35	0.65	0.00	0.00	0.00	0.09	0.91	
Team4	0.00	1.00	0.00	0.00	0.00	0.74	0.26	1.00	0.00	0.00	0.00	1.00	0.00	0.00	
Team5	0.25	0.75	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.69	0.31	
CPCs	CPC5			CPC6			CPC7		CPC8			CPC9			
Levels	L1	L2	L3	L1	L2	L3	L1	L2	L1	L2	L3	L1	L2	L3	L4
Team1	0.00	0.94	0.06	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00
Team2	0.00	0.99	0.01	0.93	0.07	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00
Team3	0.00	1.00	0.00	0.93	0.07	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00
Team4	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00
Team5	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.73	0.27	0.00	0.00	1.00	0.00	0.00

Table 5.2 Failure rate values for cognitive failure modes of each team

	O1	O2	O3	I1	I2	I3	P1	P2	E1	E2	E3	E4	E5
Team1	0.00016	0.0112	0.0112	0.025	0.00125	0.00125	0.0005	0.0005	0.000384	0.000064	0.000384	0.000384	
Team2	0.0002	0.014	0.014	0.025	0.00125	0.00125	0.00125	0.000125	0.0006	0.00006	0.0001	0.0006	0.0006
Team3	0.0004	0.028	0.028	0.025	0.00125	0.00125	0.00625	0.000625	0.0012	0.0012	0.0002	0.0012	0.0012
Team4	0.0001	0.007	0.007	0.025	0.00125	0.00125	0.00125	0.000125	0.0003	0.0003	0.00005	0.0003	0.0003
Team5	0.0002	0.014	0.014	0.025	0.00125	0.00125	0.00125	0.000125	0.0006	0.0006	0.0001	0.0006	0.0006

5.5 Robot Reliability Analysis

Robot reliability data as discussed in Section 4.6, the basic robot function failure rates are estimated as Table 4.2 shows. A communication function is added to express the human-robot communication channel failure.

In addition, the average performance metric for the crack identification algorithm used in this research is the precision at 90% and the recall at 99%. The definitions of Precision and recall are [101]:

$$Precision = \frac{tp}{tp+fp} \tag{5.4}$$

$$Recall = \frac{tp}{tp+fn} \tag{5.5}$$

Where "tp" represents the number of true positive results, "fp" represents the number of false positive results, and "fn" represents the number of false negative results.

5.6 IDDA Application

Input files for two scenarios were developed according to the system function analysis. The input files code are shown in Annex B (B.2 and B.3).

5.7 Conceptual model design

5.7.1 Agent model

Each agent is independent, can communicate with other agents, adapt to environmental changes, and influence the environmental states. An agent can be expressed as a mathematics definition as follows:

$$Agent = \{Id, A, B, R, F, G\} \tag{5.6}$$

Where Id is the index of an agent, $A = \{a_1, \cdots, a_n\}$ is a set of attributes, $B = \{b_1, \cdots, b_m\}$ is a set of behaviors, $R = \{r_1, \cdots, r_j\}$ is a set of results of related behavior, $F = \{f(x)_1, \cdots, f(x)_k\}$ is a set of the relation functions about agent perceptions and behaviors,$G = \{g(x)_1, \cdots, g(x)_l\}$ is a set of the relation functions about agent behaviors and results.

5.7.2 Phases model

The FM process has three primary phases: scaffold building, weld joints polishing, and weld joints magnetic. Each phase has three indicators, as described in **??**. And they have interactions with each other (in Fig. 5.3). The first phase, building on time, will be influenced by building stability and safety. Then, the second phase, "on time," will be affected by the first stage on time, polishing quality, and polishing safety. The polishing safety is also influenced by building stability. The second and first stages will impact the last phase on time on time, magnetic quality, and magnetic safety. In addition, the magnetic quality will be affected by the polishing quality. The magnetic safe will be influenced by building stability. This phase's logic is likewise for the HRT process, excluding the building scaffold phase.

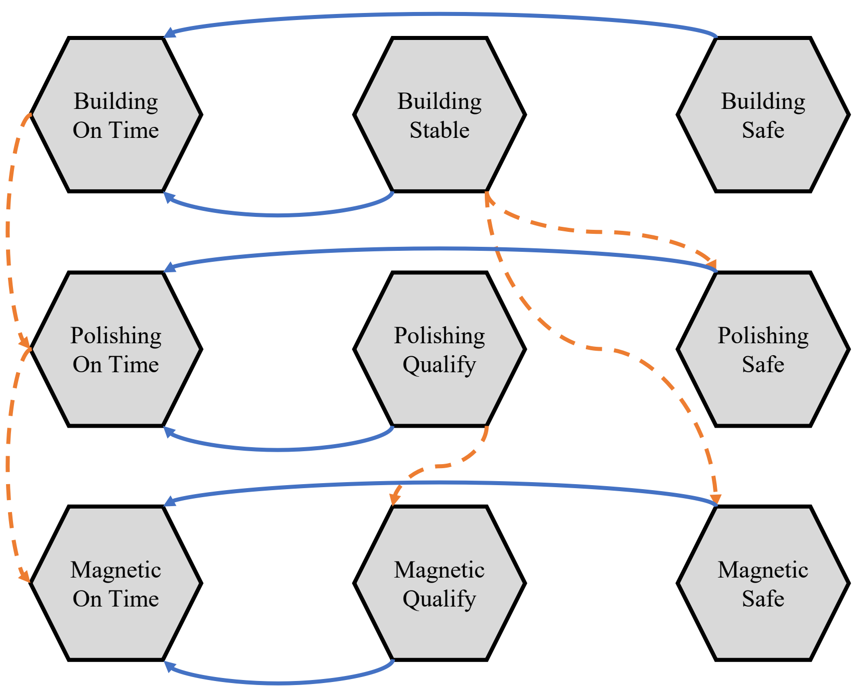


Fig. 5.3 Critical indicators of different stages and their interactions

5.7.3 Total time model

In the FM inspection process, the total time is mainly determined by the work speed of human workers. Also, the incident and recovery of the mistakes will delay the procedure.

In the HRT inspection process, the total time is influenced by the working speed, moving speed, and moving path setting. Also, the robot malfunctions and maintenance duration delay the process.

5.7.4 Miss rate model

In the FM inspection process, the polishing workers' errors may lead to low-quality polished weld joints in dynamic HEP. Then, the two polishing supervisors will randomly choose four weld joints to check, and the technician supervisors will double-check four randomly selected weld joints. Repolishing will start if any low-quality polished weld joint is found among these eight weld joints. Based on the previous work, the repolishing velocity will be three times the regular working speed. Moreover, the magnetic workers' errors in the spray and magnetic process may lead to low-quality magnetized joints in dynamic HEP. The magnetic process cooperates in a three-technician group. The technician supervisor will check the magnetic process and correct the error in dynamic HEP. Finally, if the polished weld joints are low quality, the magnetized joints are low quality, or the technician makes an error in recognizing the crack, the number of missed or false true cracks will increase. Then, the miss rate could be calculated by the formula 4.2.

In the HRT inspection process, when a polishing robot is performing the polishing work, if it is in the polishing brush failure, the iron pieces accumulate too many, or the automatic navigating laser sensor failure states and the alarms are not detected by technicians (the detection is in dynamic HEP). The polished weld joints will be low-quality. Furthermore, the robot will keep on sending alarms when in failure states. If technicians detect the alarms and the control system works well, they will stop the robot and perform maintenance and recovery. Moreover, if the automatic navigating laser sensor fails when a magnetic robot performs magnetizing work, the magnetized weld joints will be low-quality. This also could be maintenance and recovery if the control system works. If there is a real crack, if the AI image identification software fails with the lambda AI rate, the missed crack number will increase. In addition, if the AI identifies a real or false true crack, it will alarm the technicians. If they miss this alarm in the dynamic HEP, the missed crack or false true crack numbers will increase.

5.7.5 Occupational risk model

For the FM process, the construction workers must work at height in a confined space with air quality detection equipment. Besides these risks, the polishing workers must also work with dangerous tools such as polishing sander. However, these incidents could be prevented by personal protection equipment (PPE), where the failure rate of not wearing a PPE is both the initial failure rate of the worker and the supervisor. According to the literature [102], the consequence scores can be represented by different scores as Table 5.3 shows. Then, the consequence scores for incident types 1 to 4 were assigned to:

1. Incident 1 hurt by dangers tool, C=1
2. Incident 2 suffocation, C=50
3. Incident 3 fall or hit by a fall-down object, C=25.

Table 5.3 Linguistic scale for consequence

Description	Rating
Minor cuts, bruises, bumps	1
Disabling injuries	5
Extremely serious	15
Fatality	25
Multiple fatalities;	50
Catastrophe	100

5.8 Computerized Model

5.8.1 Agent hierarchy

The manual model has three kinds of agents: the clock agent, the human agent, and the joint agent. The HRT model has five types of agents: the clock agent, the human agent, the robot agent, the joint agent, and the requirement agent. The hierarchical structure is described similarly to the study [78] in this research to express the agents' relationship, as shown in Figure 5.4, three layers in this structure. The first layer is the simplified layout of the inside of the spherical tank and work site where technicians monitor the robots' work. The scaffold is simplified as a ladder to show the distance workers need to climb. In addition, this layer includes the system parameters as the model inputs and outputs.

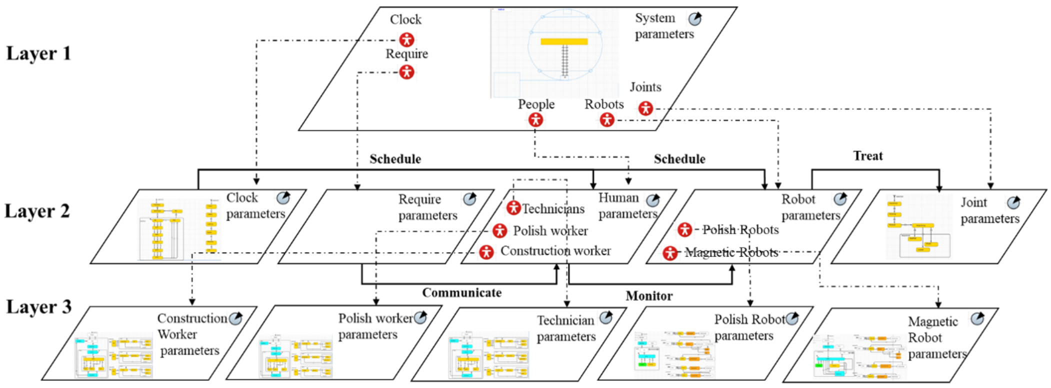


Fig. 5.4 Agent hierarchy structure

The clock agent schedules the working and rest duration, and the workers will shift in two hours as the air quality is low. The HRT scenarios set the clock as the same to compare the performance between the two systems better. Furthermore, a parallel state chart in the clock agent schedules the project phases as described in Section 5.3.2. The requirement agent is designed to record the arranged or randomly generated task or alarm for human workers in a requirement list to communicate them to a worker in the monitoring state and with attention (expressed as human reliability).

In the second layer, the internal states and transitions of the five agents are designed. The third layer includes the extended type of human agents, such as construction workers, polish workers, and technicians. In addition, this layer consists of the extended type of robot agents such as polish robots and magnetic robots.

5.8.2 Modeling the work schedule

As shown in Fig. 5.5, The state chart on the left simulates the daily work shift. While a function code is developed to get workers' states, if they have not finished the tasks or alarms in the requirement list, the work shift will be delayed, and the delayed time will be recorded. The state chart on the right simulates the project phase, shifting the work states of human workers and robots.

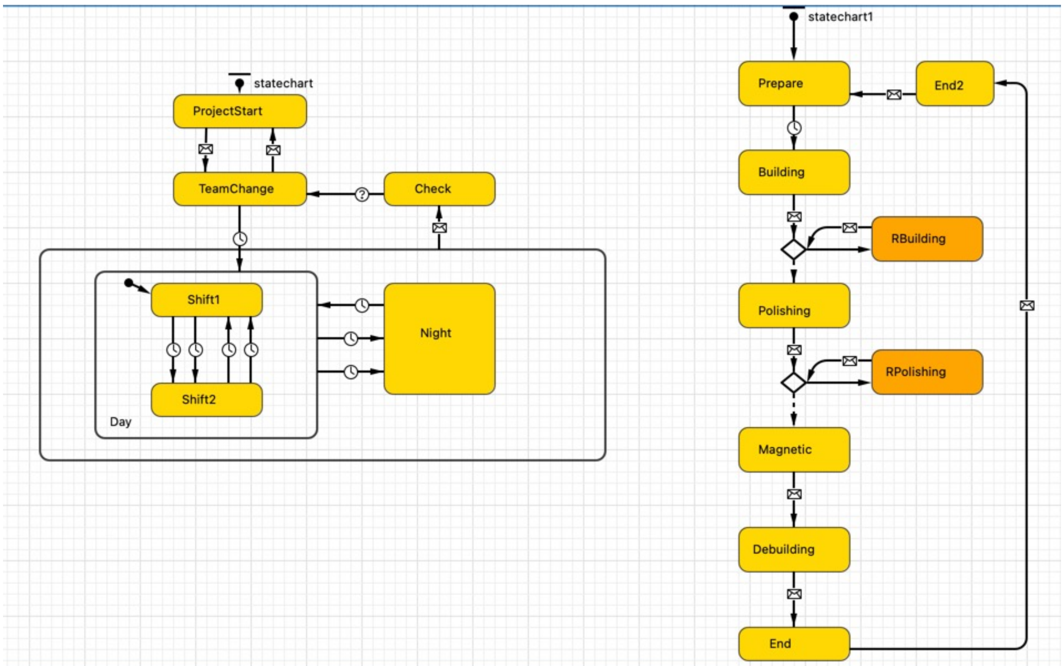


Fig. 5.5 The state chart of the clock agent

5.8.3 Modeling human workers

The common factors of human workers:

1. Dynamic human error probability. Most human reliability models, like THERP, SPAR-H, and CREAM, consider the human error rate a fixed value during one task period. However, previous studies found that the performance of humans

reduces over time due to fatigue, stress, and loss of concentration accumulated over long, continuous working time[103]. Figure 5.6 shows the causes of fatigue. Givi proposed the fatigue-recovery and learning-forget human failure rate calculation model based on the learning curve theory and tested it in the assembly industry scenarios [92]. The fatigue value could be calculated as the following formula: Equation 4.1, 4.2, and 4.3. In this research, for the FM scenario, the fatigue index is set as 0.0192, which means exhausted after six hours of work, and the recovery index is set as 0.0144, which means recovery after eight hours of reset. In contrast, for the HRT scenario, having better working conditions. The fatigue index is set as 0.096, which means exhausted after twelve hours of work, and the recovery index is also set as 0.0144.

2. human error recovery. This research considers recovery factors, including self-check, supervisor check, cross-team check, and government check.

3. behavior uncertainty. The triangular distributions with a 20 % variance in speed and check interval represent the human workers' capability differences as literature usually do[104].

The construction worker needs to perform the scaffold building work. The polish worker needs to complete the polish job. And the technician needs to perform the magnetic work. Human failure logic involves working at height, working inside a confined space, and working with sharp edges and dangerous tools.

As shown in Figure 5.6, human workers have one operation state chart. In addition, they have a failure state transition state chart. The failure state will be triggered by rate, and personal protection equipment can be a barrier. Therefore, the result will be delay or incident, which happens when humans fail to forget to wear PPE, and the supervisor does not discover this.

The technician in the HRT model needs to perform the robot setting, the robot monitoring, the robot regular maintenance, the robot response maintenance, and the magnetic crack check task.

5.8.4 Modeling the robot agent

The function failure rate estimations of the robot agent have been discussed in Section 4.6. The failure modes of polish robots are analyzed firstly as polish functions fail

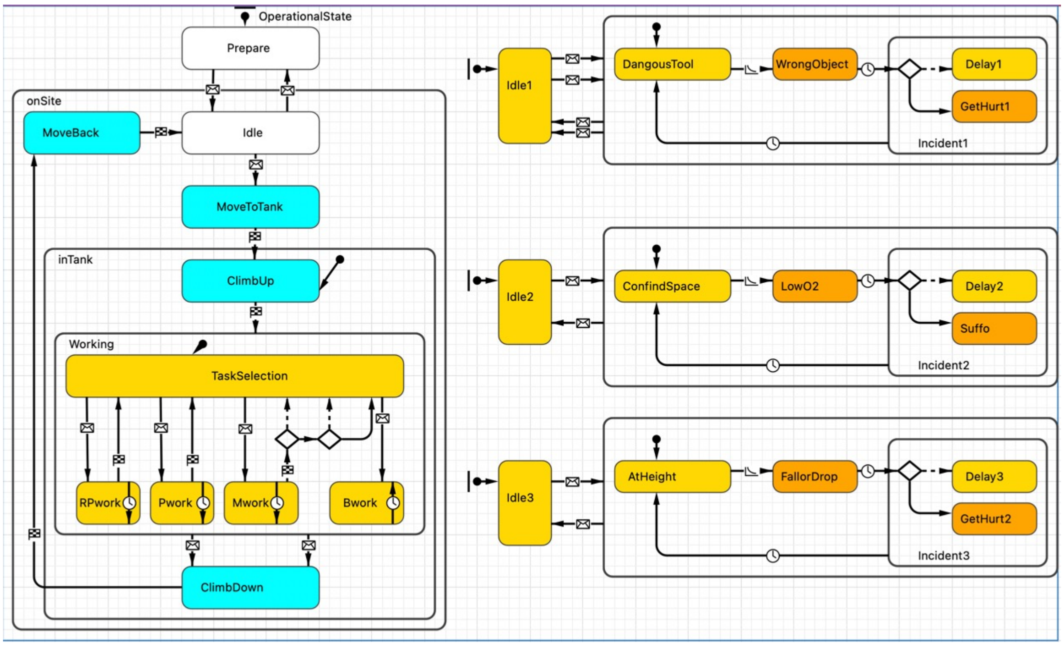


Fig. 5.6 The state chart of human worker agents

when introduced by a brush needing changing or iron pieces needing cleaning. Secondly, as move function fails introduced by end-effector fail. Control function failure is introduced by communication, a video sensor, or control system failure.

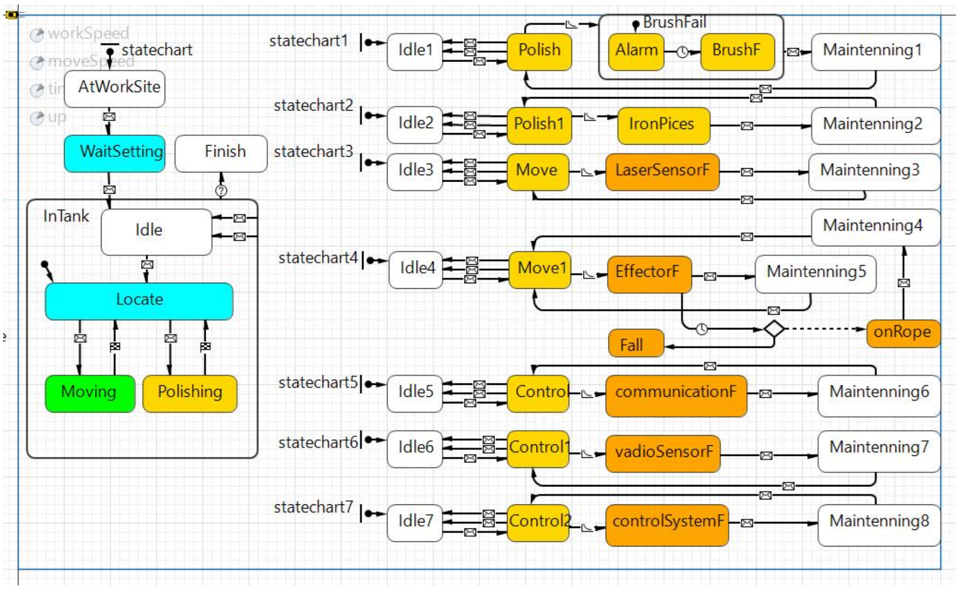


Fig. 5.7 The state chart of polishing robot agents

The state chart of polish robot is shown in Figure 5.7, and the state chart of magnetic robot is shown in Figure 5.8

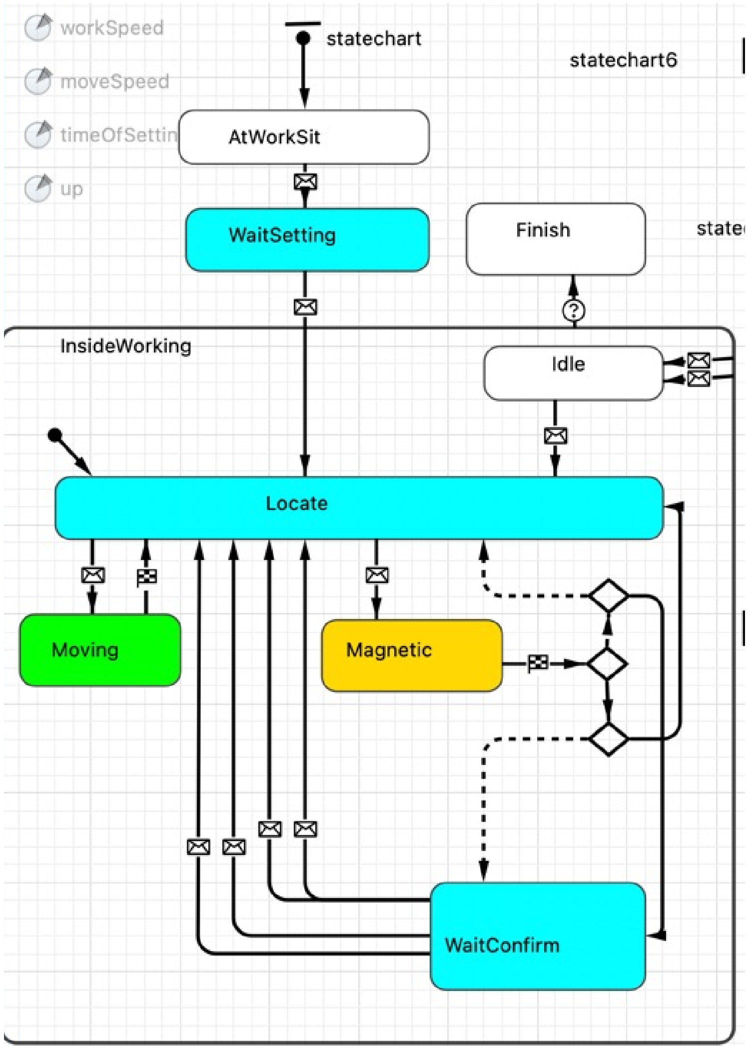


Fig. 5.8 The state chart of magnetizing robot agent

5.8.5 Modeling the human-robot interaction

The human-robot interaction in ABMS could be represented by message exchange. As Figure 5.6 shows, when the magnetic robot is in the state "WaitSetting," it will send a message "waiting" to the human technician, which will trigger the human technician to move to the robot. When the technician arrives at the robot and finishes the setting, he can send the message "settled." to the robot, which will trigger the robot to begin working. Also, when the magnetic robot identifies a crack, it will send the alarm to the human technician and wait for the human technician's decision. The human technician's decision will also be sent to the magnetic robot by message.

5.9 Assumptions and Experiment design

The assumptions for the model building include the following:

1. There is no more than one crack on one weld joint. The total number of no cracks is set as a fixed number of 24. This setting is for the convenience of simulation and comparison of the wrong result rate of the inspection process at the same starting baseline.

2. In both systems, The shift schedule is set for two hours of work and rest for two groups, each group working up to 8 hours per day.

The primary settings for the HRT system model are two polishing robots, two magnetic robots, and four technicians. According to the real operation requirements, the FM system model resource settings are 18 construction workers, 6 Polish workers, and 6 technicians.

Anylogic software is used to perform the ABMS Monte Carlo and sensitivity experiments. All the human behavior parameters are randomly generated from the pre-set distributions. The Monte Carlo experiments for the FM and HRT operations were run 5000 times with random seeds and every 10 data as a group to calculate the mean value for total time and wrong result rate according to Equations 5.1and 5.2. The mean occupational risk value was calculated based on extra data from Monte Carlo experiments in the FM scenario, including 200000 data, and the bootstrapping method was employed for every 10000 sampled data in a group to run 1000 times to catch better the occurrence probability of an occupational accident.

For the sensitivity test, the experiment setting for each parameter is shown in Table 5.4. The sensitivity experiment for each step value of one parameter ran 500 times to catch the random behavior, and every 10 data as a group was used to calculate the mean results.

Table 5.4 The parameters setting for the sensitivity test

Parameter	Min	Max	Step
HEP_{base}	0.001	0.01	0.001
HEP_{base} in the HRT scenario	0.1	0.5	0.05
$\Lambda_{fatigue}$ in the FM scenario	0.0144	0.028	0.002
$\Lambda_{fatigue}$ in the HRT scenario	0.0096	0.0144	0.002
$AI_{missrate}$	0.01	0.25	0.02
$AI_{falsetruerate}$	0.05	0.4	0.05

Chapter 6

Results

Qualitative and quantitative results were gained after the integrated framework was applied to the on-site case study. Related work has been published in Safety Science, entitled"Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations" [26]. Another paper in Advanced Engineering Informatics entitled " Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis" is under review.

6.1 System performance qualitative analysis results based on task analysis

Generally, there are several communication modes (direct physical interaction, remote contactless interaction, tel-operation, and message exchange) and interaction modes (coexistence, synchronized, cooperation, collaboration) [105], according to the integrated task analysis, the human-robot communication modes and interaction modes are summarized in Table 6.1 and Table 6.2[26]

Table 6.1 The human-robot communication modes

Communication mode	Scenario
Direct physical interaction	The technician removes iron pieces from the polishing robot
Remote contactless interaction	Robots send an alarm to alert the technician The magnetic robot waits for confirmation of the identified crack The technician monitors robots
Tel-operation	The technician controls the robots to the right paths remotely
Message exchange	The technician set the robot parameters

Table 6.2 The human-robot interaction modes

Interaction mode	Scenario
Synchronized (in different places)	The magnetic robot, and the technician perform the crack identification task. The robot first identifies the crack and stops to send an alarm to the technician. Then, the technician detects the alarm and confirms or denies the crack identification decision.
Collaboration	The technician removes iron pieces from the polishing robot. The technician and robot perform the polishing task together.

Goodrich & Schultz propose the concept of dynamic interaction as a characterization that incorporates all five dimensions HRI designers can affect, namely autonomy, how information is exchanged, team structure, learning and training of the humans and robots involved, and the shape of the task [106]. The changes HRI brings to the system in this case are summarized in Table 6.3.

Table 6.3 The human-robot interaction changes

	Team structure	Communication type	Human role	Task demand Cognitive function
FM	5 teams	Multi-team communication.Inter-team communication;	Single role as operator or supervisor	More executions
HRT	2 teams	Inter-team communication.Human-robot communication;	Multi roles as both operator and supervisor	More Interpretation and plan

6.2 System performance quantitative analysis based on IDDA

According to the literature [102], different scores can represent the consequence scores: Minor cuts, bruises, and bumps at 1. Disabling injuries at 5. Extremely serious at 15, Fatality at 25, Multiple fatalities at 50, and Catastrophe at 100[26]. In addition, the time delay consequence scores were set as the corresponding delay time duration with the unit as one day. With this definition of consequence scores and the IDDA input files, the results of all the event trees for each scenario can be generated with the risk. An example of an IDDA input file and output is shown in Annex B (Table B.1 and Table B.2).

6.2.1 The system performance in the FM scenario

Fig. 6.1 shows the values of occupational incident risk in the FM scenario[26]. Task 3(Manual polishing) has the highest probability of occurrence at 8.00×10^{-4} and risk at 1.72×10^{-3}, followed by Task 4 (Manual magnetic and crack identification) [26]. Among all the subtasks, the object or personnel that fell from high has the highest risk at 1.43×10^{-3} [26]. A spark from polish that hurt the operator's eye at 9.61×10^{-4} [26].

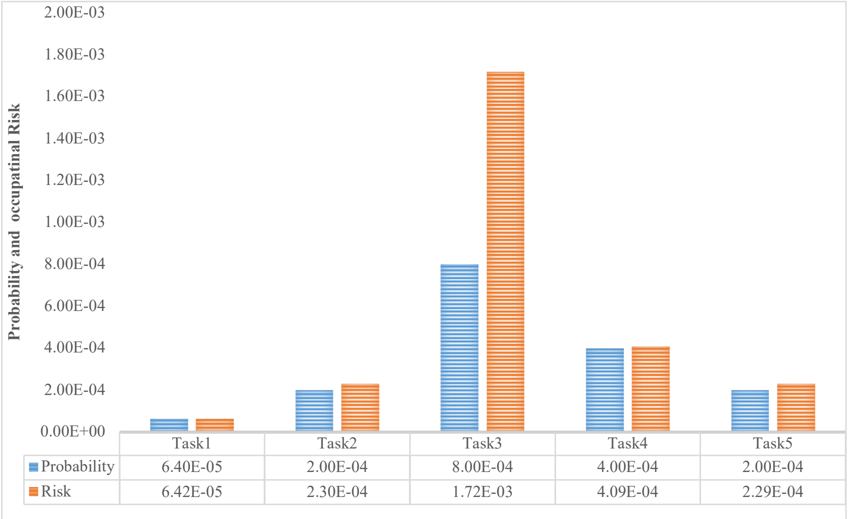


Fig. 6.1 The occupational incident probability and risk for FM main tasks

Fig. 6.2 shows the values of time delay risk in the FM scenario [26]. Task 2(Scaffold building) has the highest risk at 4.30×10^{-3}, with the third highest probability at 8.00×10^{-4} [26]. Among subtasks, fixing the unstable scaffold has the highest risk at 3.71×10^{-3} [26]. Repolishing the weld joints operation risk is at 1.22×10^{-3} [26]. Moreover, the third risky delay subtask is reventilating the spherical tank with risk at 1.02×10^{-3}[26].

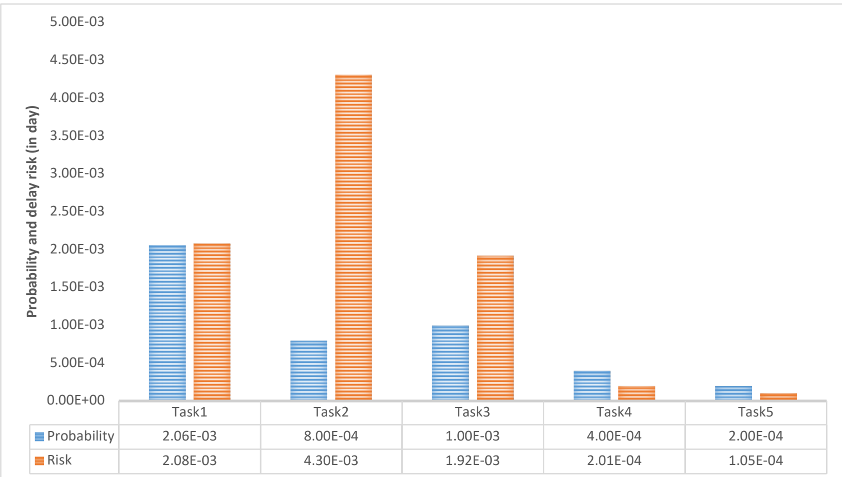


Fig. 6.2 The delay probability and risk for FM main tasks

Only Task 3 (Manual polishing) and Task 4 (Manual magnetic and crack identification) may lead to the wrong result of the crack identification, with the total probability at 1.53×10^{-2}[26]. The most contributing subtask is humans identifying the crack wrongly with probability at 1.27×10^{-2}.

6.2.2 The system performance of the HRT scenario

Fig. 6.3 shows the values of occupational incident risk in the HRT scenario. Task 2 (HRT polishing) and Task 3 (HRT magnetic and crack identification) have the same failure rate at 2.00×10^{-4}, followed by Task 1, which is the same in the FM scenario[26]. Falling when climbing the ladder outside the tank has the highest risk at 2.24×10^{-4}. Getting shocked when connecting the robot to electricity has the second highest risk at 1.96×10^{-4}. [26].

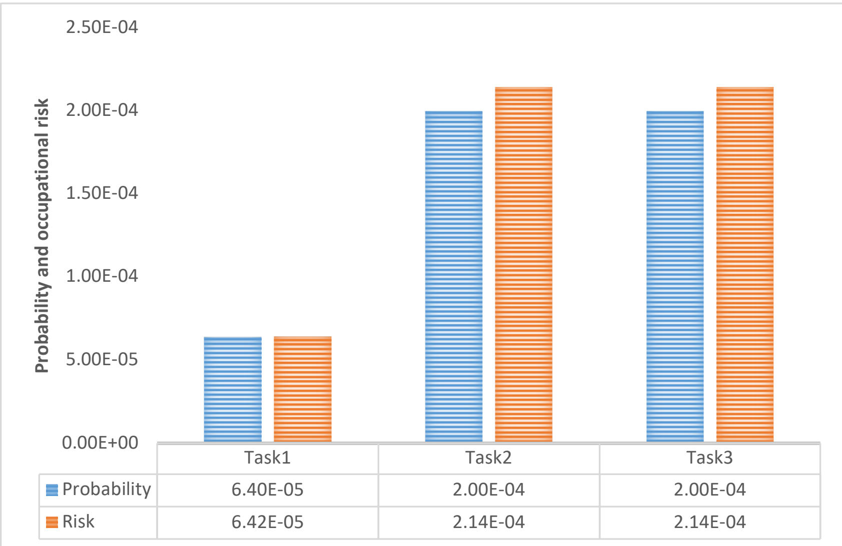


Fig. 6.3 The occupational incident probability and risk for HRT main tasks

Fig. 6.4 shows the values of time delay risk in the HRT scenario. Task 1, the same as the FM scenario, has the highest risk, with the lowest probability at 2.06×10^{-3} [26]. Task 2 (HRT polishing) and Task 3 (HRT magnetic and crack identification) have a slightly lower risk at 1.58×10^{-3}. The failure rate of the HRT system to get the wrong result of crack identification was 1.58×10^{-3}[26]. Not cleaning the iron pieces in time contributes most to the delay and the risk of wrong results in HRT operations[26]. This is followed by the operator refusing the robot's decision, which has correctly identified a crack with a probability at 1.18×10^{-2}.

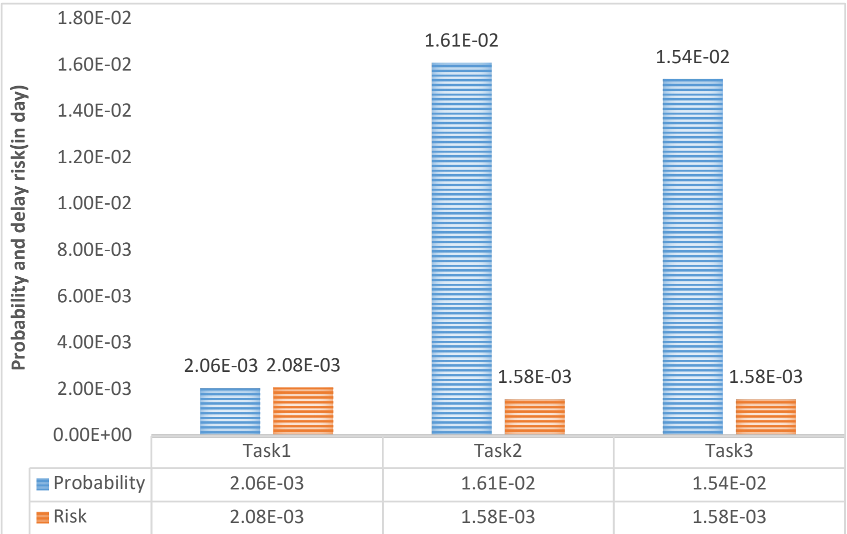
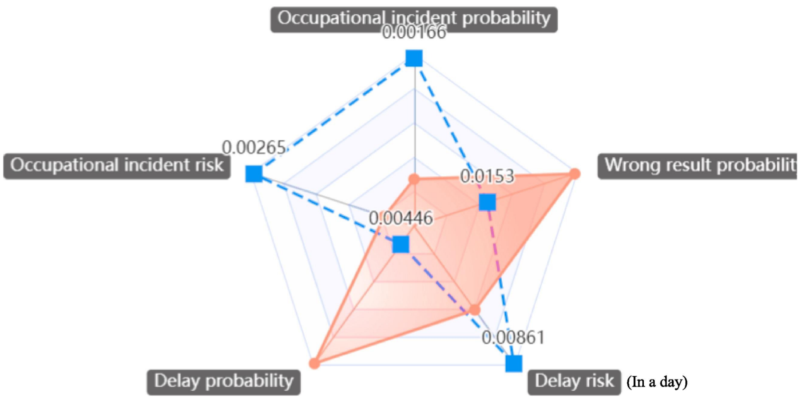
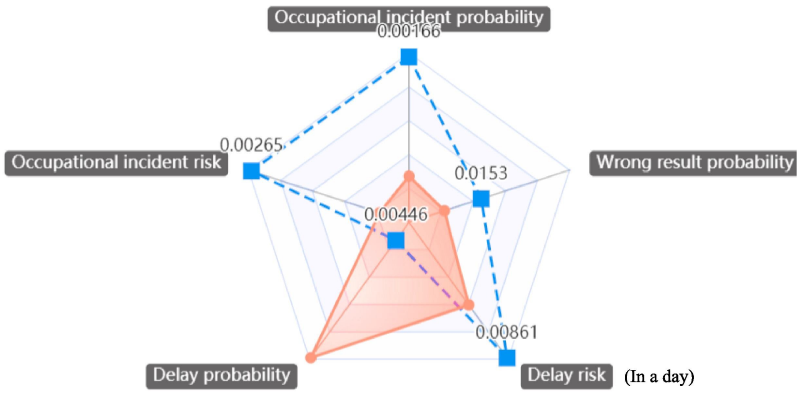


Fig. 6.4 The delay probability and risk for HRT main tasks



(a) System Performance comparison before optimization



(b) System Performance comparison after optimization

Fig. 6.5 Overall comparison of the FM and HRT

6.2.3 The system performance comparison of two scenarios

Fig. 6.5 shows the system performance comparison of the FM and HRT scenarios. The HRT system shows the lower occupational incident probability and risk at the values 4.64×10^{-4} and 4.93×10^{-4} [26]. This risk is nearly five times less than the FM system. Also, the HRT system has less delay risk than the FM system at 5.24×10^{-3}, although with a higher delay probability 3.36×10^{-2} [26]. This difference means the HRT system may have more delay scenarios, but the overall delay time is less than the FM system. However, the probability of getting the wrong result of the HRT system was higher, at 3.35×10^{-2} [26]. After checking the wrong result leading subtasks, the optimization measurement could be added to a double check for the identified cracks to reduce the false true cracks. Utilizing the IDDA method, it is possible to reduce the probability of getting the wrong result for the HRT system to 7.52×10^{-3} [26].

6.3 System performance based on ABMS

6.3.1 HRT operations performance assessment and comparison

After the Shapiro-Wilk test [107], the total time and wrong result rate results are not from a normal distribution with p-values smaller than 0.001. Their density distributions are shown in Fig. 6.6 A and B; the results demonstrated that the HRT system performance indicators had significantly lower variance in total time and wrong result rate, indicating the HRT had higher system reliability. The two-sample Wilcoxon rank-sum (Mann-Whitney) test [108] is selected to compare the FM and HRT operations' data. Results show that the HRT process could reduce the total time by more than half. The median value of mean total time changed from 432.25 to 181.60 (Fig. 6.6 (C)). As shown in Fig. 6.6 (D), the median value of the wrong result rate decreases from 0.04 to 0.01. This result is based on the image identification miss rate of 0.01 and the false true rate of 0.1, which is an ideal value for the AI image identification algorithm.

For safety, the mean occupational safety risk in the FM process is 6.23×10^{-3} with a probability of 2.50×10^{-4} in the FM operations. In contrast, the value for the HRT process is nearly zero. In addition, the mean in-tank time of the construction

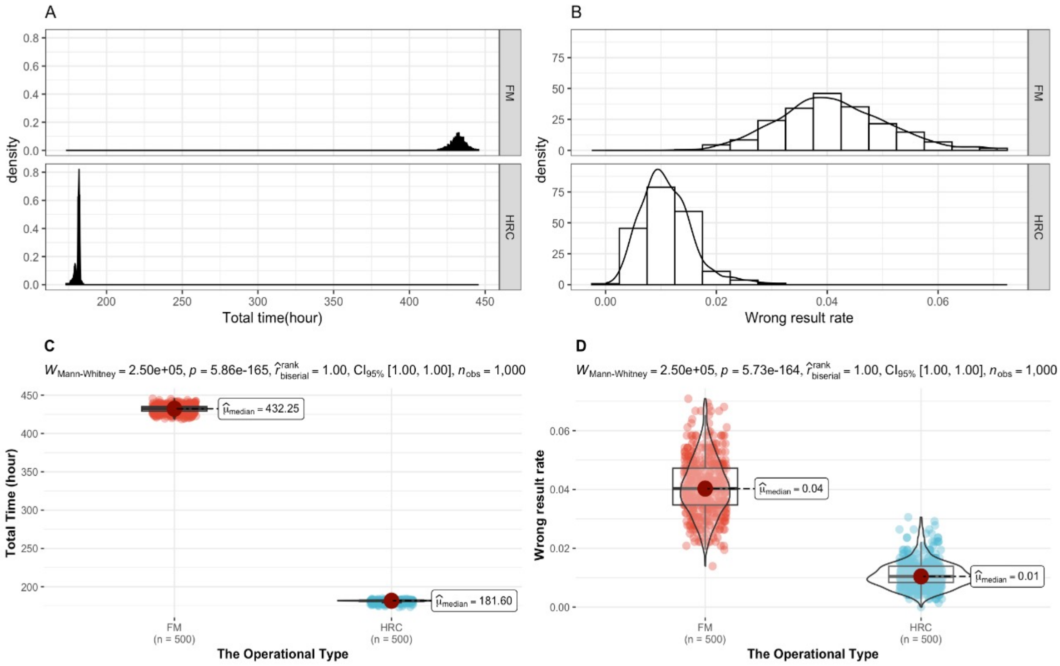


Fig. 6.6 System performance comparison between two operational types

workers is 55.74, the mean in-tank time of the polish workers is 42.11, and the mean in-tank time of the technicians is 47.32. In contrast, the technician's in-tank time in the HRT system is less than one hour. In this way, the HRT process could reduce contact with occupation hazards by a great deal.

6.3.2 HRT scenario sensitivity analysis

The mean wrong result rate of the HRT process is significantly sensitive to the miss rate of image identification algorithms ($AI_{missrate}$=1-recall). As the $AI_{missrate}$ increases, the mean value of the wrong result rate tends to increase as well and spread to a broader range, as shown in Fig. 6.7 B and D. While the mean total time is not sensitive to $AI_{missrate}$, this indicator tends to decrease, as shown in Fig. 6.7 A and C. These results could be explained by the HRT procedures to identify the cracks. The magnetic robot performed the first step with AI image identification algorithms; if this step misses a crack, there is no way to recover this mistake. If this step gives a false true alarm, then the operator still has a chance to recover it with first and second checks.

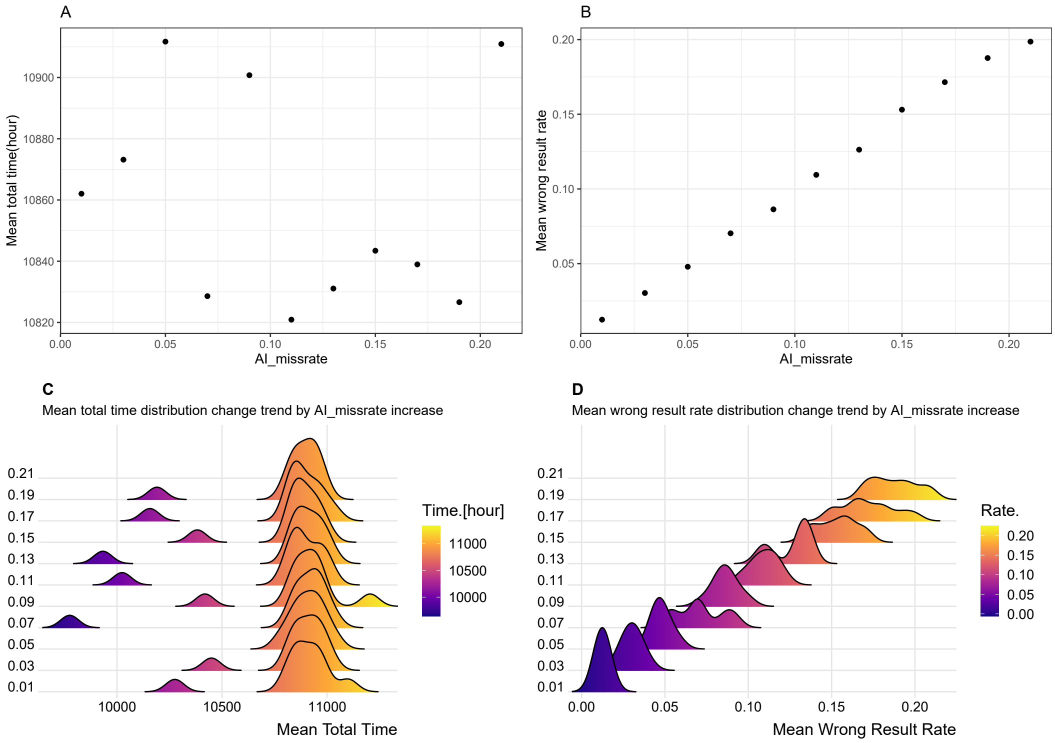


Fig. 6.7 System performance sensitivity to $AI_{missrate}$ in HRT operations

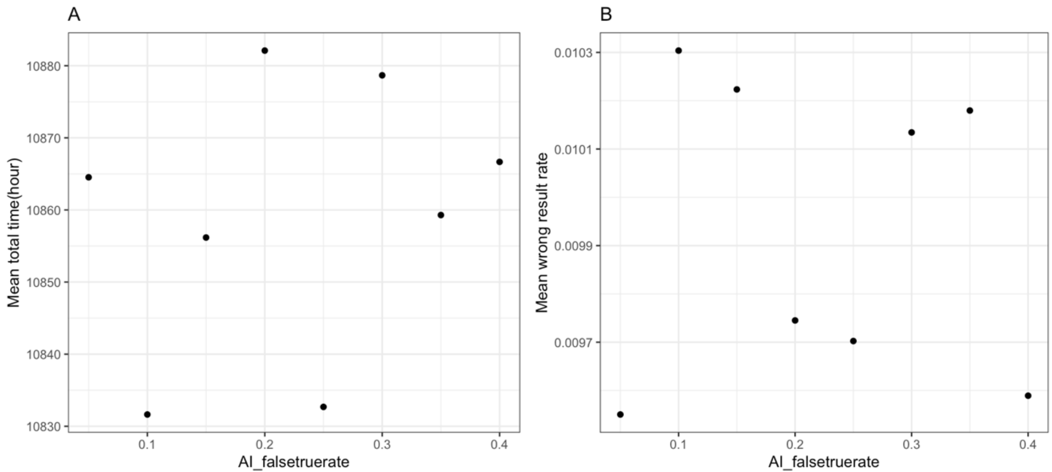


Fig. 6.8 System performance sensitivity to $AI_{falsetruerate}$ in HRT operations

As shown in Fig. 6.8 A and B, the HRT system performance is not sensitive to the $AI_{falsetruerate}$. This situation could be explained by the false true crack, which could be recovered by human technicians by the first check.

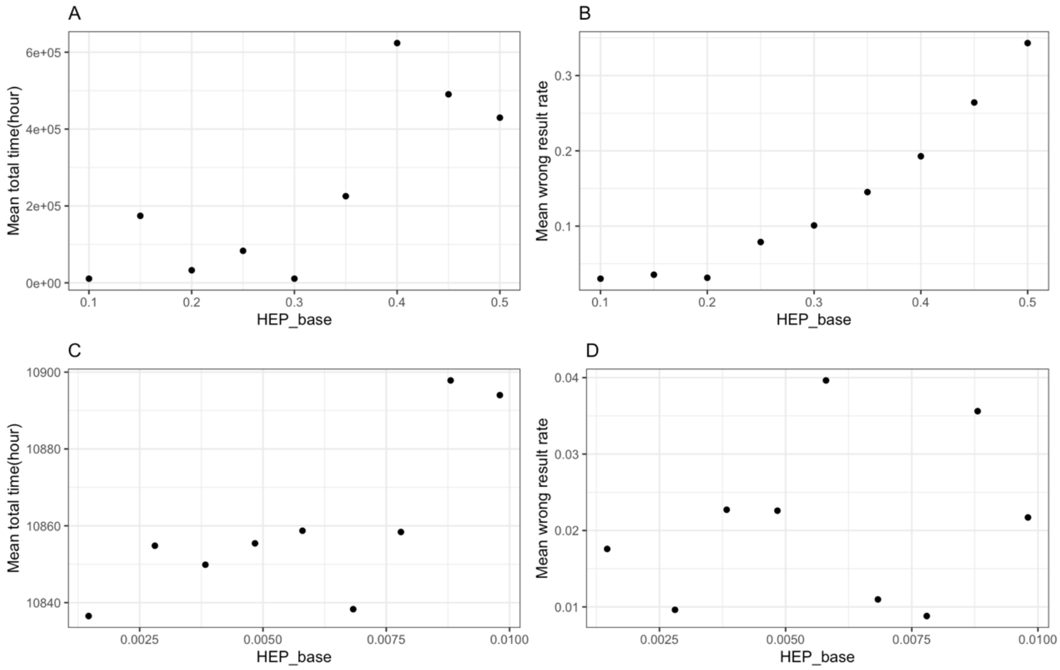


Fig. 6.9 System performance sensitivity to HEP_{base} in HRT operations

As shown in Fig. 6.9 C and D, when the HEP_{base} is lower than 0.1, the performance of the HRT process is not sensitive to the HEP_{base}, the same situation as the test of $\lambda_{fatigue}$ sensitivity. However, as shown in Fig. 6.9 B, when the HEP_{base} is higher than 0.1, the mean wrong result rate of the HRT process is sensitive to the HEP_{base}. This change could be explained when the human technician wrongly changes the right decision from the AI algorithm with a more significant error probability than the $AI_{falsetruerate}$, being a major influencing factor to the wrong result rate.

6.3.3 FM scenario sensitivity analysis

As Fig. 6.10 and Fig. 6.11 are shown, the performance of the FM process is sensitive to the fatigue index and HEP_{base}. The mean values of performance indicators increase along with the $\lambda_{fatigue}$ increase, as Fig. 6.10 and Fig. 6.11 A and B show. Also, the variance of wrong result rate increases, as shown in Fig. 6.10 and Fig. 6.11

D. This means the increase of HEP_{base} or $\lambda_{fatigue}$ will introduce more difficulty in process quality control. This influence could be attributed to the FM process performed by human workers directly.

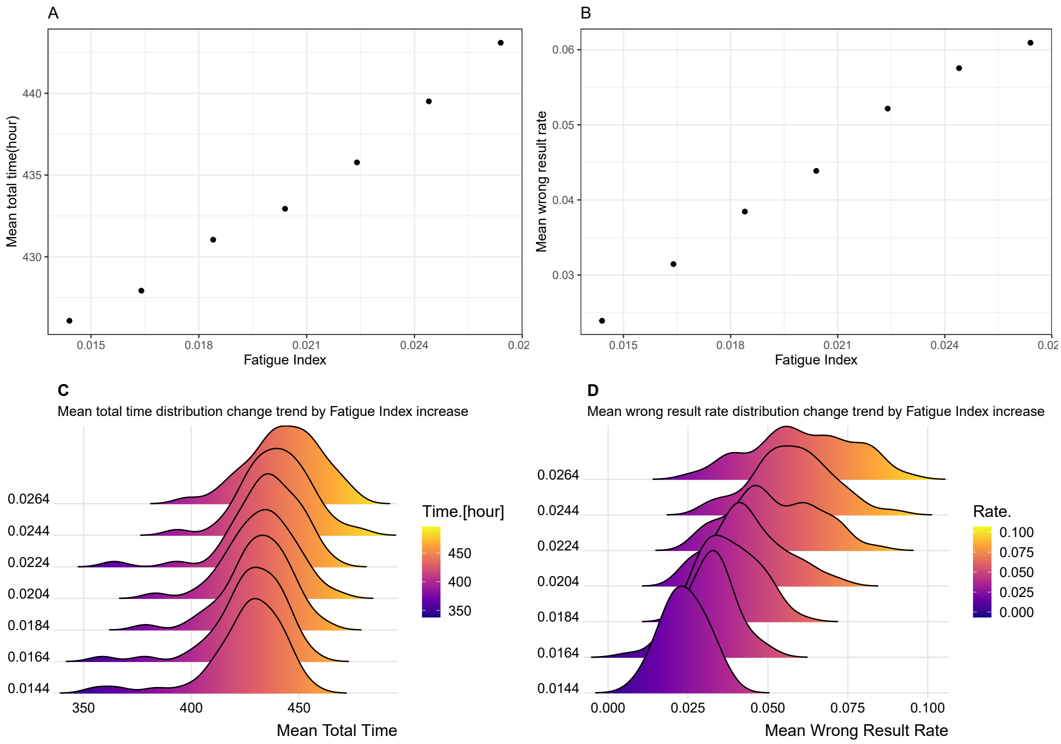


Fig. 6.10 The system performance change trend by $\lambda_{fatigue}$ increase

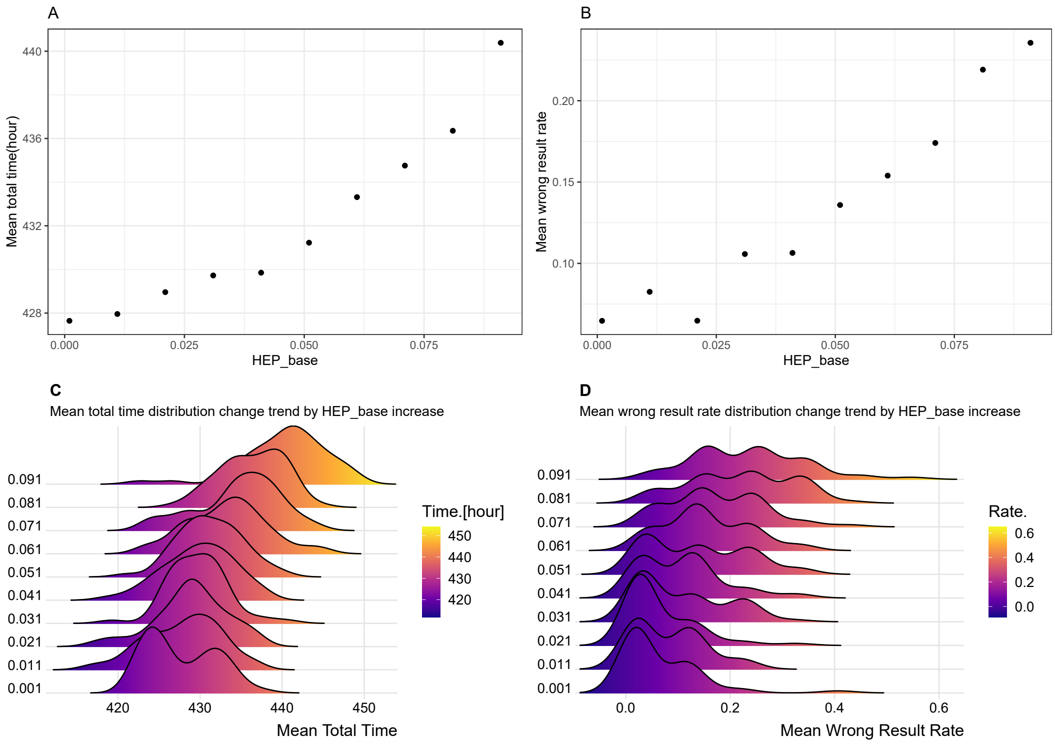


Fig. 6.11 The system performance change trend by HEP_{base} increase

Chapter 7

Discussion

There is a trend in Industry 4.0 to employ a robot to assist human operators as a teammate. This transition also introduces new changes to system performance. Very few research studies have focused on comparing system performance between FM and HRT operations. The HRT system is becoming more complex, and the traditional way of risk analysis, considering only the HRA or equipment reliability, is no longer satisfying. This thesis develops an integrated framework for complex system performance assessment, which considers HOFs, technical factors, and the interactions among them simultaneously. The case study shows that this framework could generate knowledge about multi-dimensional performance and its critical influencing factors for both FM and HRT systems. Related work has been published in Safety Science, entitled"Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations" [26]. Another paper in Advanced Engineering Informatics entitled " Evolving Process Maintenance through Human-Robot Collaboration: An Agent-based System Performance Analysis" is under review.

7.1 Comparison with previous research

The two scenarios of the pressure spherical tank inspection case show that the proposed integrated framework can do qualitative and quantitative analysis for the complex socio-technical system performance[26]. Compared to the HRT design framework proposed by Borges et al.(2021), it was mainly focused on applying scale tools in ergonomics [15]. The framework in this research emphasizes the integrated task analysis, considering the cognitive functions and the function interdependence. In this way, the case study shows a more detailed logical foundation for the quantitative analysis[26]. Based on qualitative analysis results, as shown in Table 6.1, HRT's most essential communication mode is human-robot message exchange. In this way, the performance of the HRT system is deeply impacted by human's rational trust in robots [109]. Wang et al. designed a simulated experiment to let a robot give its algorithm accuracy values to the human teammate through natural language expressing to add the human decision-making on whether to trust a robot's judgment [110]. This expression method could be an alternative to enhance information transparency in the process of crack identification. That is, the robot gives precision and recall simultaneously with the crack identification decisions. For example, the robot gives the crack judgment and the message, "I believe this is a crack, but I may be false true at 10%". This dialogue could aid the human teammate in better informative decision-making.

Based on the quantitative analysis of the FM system risk assessment, falls from height are the scenarios that contribute the most to occupational accidents in the inspection process. This is consistent with previous literature [111]. The overall human failure rate of scaffolding work was estimated to be 4.30×10^{-3} [112]. In this research, the total failure rate of Task 2 (Scaffold building) is at 4.73×10^{-3} (combining the time delay and occupational accident probability). Our results are higher for this research considering the inside tank hazards. This similarity of results validates the effectiveness of this framework.

The comparison of quantitative results showed that the HRT scenario reduced occupational risk. This could be explained by robots working in an extreme environment instead of humans. However, the downside of this change may lead to the loss of some precision of the crack identification[26]. The task that most contributes to the overall failure probability of the wrong result in the HRT operation scenario is "detecting the iron pieces accumulated in the polish robot." In the earlier HRT operation design, the detection function relied on estimating time intervals and identifying radiation from the robot body. The improved suggestion for the HRT system design includes adding a feature that enables the robot to self-identify accumulated iron pieces and send an alarm with a distinct sound and a colored icon. The second contributing subtask is the operator refusing a decision on the part of the robot, which has been correctly identified as a crack by the robot[26]. This risk could be reduced by adding a second check by another technician[26]. Under this circumstance, the overall failure probability of the wrong result in the HRT operation scenario would be reduced to about 7.52×10^{-3}, nearly half of that in the FM operation scenario, which will be superior to the FM system in nearly all aspects, other than the delay probability as shown in Fig. 6.5 This is because the inspection robots are still in the testing stage which will have more failure scenario than the mature ones according to the bathtub cure [113]. Therefore, the HRT system promises better performance in all aspects in the long term. Results show that the fatigue index and HEP_{base} significantly influence FM system performance. The fatigue control measurement could be better arranged in the working schedule.

7.2 Novel discoveries

There is already some literature about the automatic crack detection model training methods development, as Table 7.1 shows.

Table 7.1 Automatic crack detection model performance in Fluorescent MPI

Learning Method	Data Source	Test dataset	Precision	Recall	Source
Two-stage CNN	Steering knuckles	10 images augment to 500 images	96.3%	90%	[99]
Two-stage MobileNetV3-CA	Bearing rings	100 images	91.7%	96.5%	[114]
ResNet50	Titanium alloy	10 plates, length >0.7mm	-	96%	[115]
Random Forest	POD test plates	18 images, 4617 windows	-	76%	[116]

It is known that there is a trade-off between Precision and Recall when learning a classification automatic detection model, especially because of the scarcity of real tank crack image data. Still, there is no recommendation about how to set up the trade-off strategy. In this research, the recall is set to 0.99, and the precision is set to 0.90. As it is validated in the sensitivity test in the HRT group, the primary critical performance for a magnetic robot for the first stage is a low miss rate (also a higher recall rate). In this way, at the first stage, human technicians could recover false positive results with less effort when looking through all the images. The human technicians' recovery operation could be a golden standard for training the model. In the long term, checking the automatically marked result will be more convenient than an omitted mistake. Look at Table 7.1 that the researchers are training the model to balance the precision and recall simultaneously, while we recommend prioritizing the recall. In this way, the robots can help reduce the workloads of human technicians and quickly improve their precision in identifying the image of the crack as a reinforcement learning process, which could work independently sometime later.

For the human operators, if the HEP_{base} is higher than 0.1, the HEP_{base} will increase the mean wrong result rate as it grows. Therefore, it is still essential in the HRT system to manage the organizational factors influencing human behavior, such as training and procedures.

In addition, a dynamic simulation model was also built as a tool to simulate the total project time for a fully manual and HRT system. This can predict the project time before carrying out the real work and help the decision-maker estimate the investment cost and resource allocation strategies.

7.3 Limitations

Certain limitations need acknowledgment: the present study relied upon robot reliability data sourced from regulatory mandates and literature reviews. It is essential to recognize that such data origins may engender a conservative estimate compared to real-world industrial implementations. Second, the proposed model is restricted to a small environment, such as the spherical tank and factory work field, ignoring the environment's random influences, such as the electricity shutdown and extreme weather, which could be included in further research. In addition, the initial attempt to set the weights of HEP_{base} and the Fatigue value in Equation 4.7 is to set them equal in this research.

Chapter 8

Conclusions

8.1 Summary of the Research

In this research, a risk-based integrated framework to evaluate system performance in complex socio-technical systems is proposed and tested in a case study in pressurized tank inspection in the FM and HRT operations to aid safety and quality management decision-making.

The rationale of this approach is founded on the complex system theory, with the characters as multiple elements, emergence at the system level, nonlinear relationship, and adaptation. Following this line of reasoning, the integrated approach is built in a holistic perspective, considering HOFs and technical factors and the interdependence between elements in terms of logical constraints or message communication. In the following, the findings for each research question are summarized:

- RQ1. What are the most critical HOFs contributing to accidents in the process industry? In Chapter 2, "4W"s information on accident reports from the eMARS dataset was investigated using NLP-based data analysis. The contributions relationship between HOFs and consequences were tested through the chi-squire method. In terms of human failure mechanisms, through an MMHC machine learning method, a kernel Bayesian Network of five basic organizational factors- design, risk assessment, training, procedure, and supervision-was learned and tested with sensitivity and predicted accuracy. Therefore, these five are the most critical HOFs contributing to accidents in the process industry. As a deeper analysis, this Bayesian Network could function as the kernel part for further research in a more complete Bayesian Network building when more detailed data is available. After the literature review in Chapter 3, the CREAM, which contains these five factors, could better represent the HRA after being extended.

- RQ2. What is the integrated system performance evaluation framework? In Chapter 4, based on the principles summarized from literature and author knowledge, an integrated framework is proposed. Firstly, the top-down qualitative phase is designed to capture the system-level emergence characters as the system goals to direct the integrated task analysis considering subtask interdependency. The outputs of this qualitative phase are system performance metrics and integrated task lists with logical interdependency. The extended CREAM method was employed to estimate human reliability data by utilizing these outputs as inputs for the qualitative phase. In contrast, data on equipment and robot reliability could be gained from the regulation and literature. Employing a bottom-up strategy in the quantitative phase. The IDDA method is selected to investigate the abstract risk of the system, with the aid of logical contain to represent the interdependency. Then, the ABMS is chosen to model the more detailed system performance as a complement and validation of the results from IDDA.

- RQ3. What are the critical differences between HRT and FM system performance? In Chapter 5, the case study of the pressurized tank inspection in FM and HRT operations, the HRT system showed superior performance in terms of total time, wrong result rate, and occupational risk when human technicians shift every two hours, and work scheduled within 8 hours a day and perform double-check the magnetic reported cracks.

- RQ4. What are the primary parameters influencing the performance of HRT and FM systems? The results in Chapter 6 show that the fatigue index and HEP_base significantly influence FM system performance. The fatigue control measurement could be better arranged in the working schedule. At the same time, the $AI\_missrate$ image is a critical influence factor to the HRT system reliability. However, the HRT system was very sensitive to the $AI_{massrate}$. This recommends the automatic image classification model training priority recall over precision and employs an enforced learning structure, then improves precision as the human technicians teach them in real operations. In addition, if the *HEP_base* is higher than 0.1, the *HEP_base* will increase the mean wrong result rate as it grows higher. Therefore, it is still essential in the HRT system to manage the organizational factors influencing human behavior, such as training and procedures.

Once again, system thinking and a holistic approach are the paradigms employed in this research to capture the complex, dynamic, and uncertain elements and system states. This is an important premise when applying the framework proposed in this thesis to other real cases or further research.

8.2 Strengths of the Research Work

The framework proposed in this work is built holistically upon complex system theory. It is advanced with the state of the arts in using more realistic ways to represent the HOFs. It considers the interdependency between system elements. Moreover, it is validated by step-by-step analysis of real maintenance scenarios.

Firstly, to represent the HOFs more realistically, the CREAM method is extended with the Dempster-Shafer theory to reduce subjective bias, and fatigue-recovery functions are employed to achieve dynamic attributes representing human behavior. Secondly, to represent the interdependence of the system element, the IDDA method is employed to represent the interdependence of the subtasks in a logically constrained way. Moreover, the ABMS method models the HRT communications by message change mechanism. Hence, the proposed framework becomes more complete than the traditional approach. Also, the on-site applications show that the proposed framework can provide both qualitative and quantitative results of HRT system performance. The applied ABMS explicated in the analysis can be used to investigate the system performance influencing factors and their sensitivities.

Therefore, the proposed framework effectively guides integrated risk and performance assessment for socio-tech complex systems like HRT systems. The qualitative results describe the changes in team structure and communication modes. The quantitative results could provide knowledge about the system performance evolving from FM to HRT in process maintenance activities, the key influencing factors, and some critical boundary values were identified. For the practical parties, we suggest that quality control and training of human operators are also needed in the HRT process.

8.3 Further work

Future research could be carried out in the following directions. First, the robot reliability parameters in this research are mainly set according to policy requirements. More actual data could be integrated into the model with the inspection robot application in practice. Second, the proposed model is restricted to a small environment, such as the spherical tank and factory work field. Furthermore, the environment's random influences, such as the electricity shutdown and extreme weather, could be included in further research. In addition, the initial attempt to set the weights of HEP_base and the Fatigue value in Equation 4.7 is set equal in this research. Further work could explore optimal weights to balance the influential effect of the HEP_base and Fatigue value according to the dynamic HEP and corresponding data generated in the models. Moreover, other practical applications of the models could be extended with optimized algorithms for the human work schedule and robot work path.

References

[1] Shuo Yang and Micaela Demichela. A data-driven narratives skeleton pattern recognition from accident reports dataset for human-and-organizational-factors analysis. *Journal of Loss Prevention in the Process Industries*, 83:105047, July 2023.

[2] Weisheng Lu, Junjie Chen, Yonglin Fu, Yipeng Pan, and Frank Ato Ghansah. Digital twin-enabled human-robot collaborative teaming towards sustainable and healthy built environments. *Journal of Cleaner Production*, 412:137412, August 2023.

[3] Tom P. Huck, Nadine Münch, Luisa Hornung, Christoph Ledermann, and Christian Wurll. Risk assessment tools for industrial human-robot collaboration: Novel approaches and practical needs. *Safety Science*, 141:105288, September 2021.

[4] Salvina Murè, Lorenzo Comberti, and Micaela Demichela. How harsh work environments affect the occupational accident phenomenology? Risk assessment and decision making optimisation. *Safety Science*, 95:159–170, June 2017.

[5] European Standard. Maintenance — Maintenance terminology, November 2017.

[6] Shuaiqi Yuan, Genserik Reniers, Ming Yang, and Yiping Bai. Cost-effective maintenance of safety and security barriers in the chemical process industries via genetic algorithm. *Process Safety and Environmental Protection*, 170:356–371, February 2023.

[7] A.R. Hale, B.H.J. Heming, K. Smit, F.G.Th. Rodenburg, and N.D. Van Leeuwen. Evaluating safety in the management of maintenance activities in the chemical process industry. *Safety Science*, 28(1):21–44, February 1998.

[8] Peter Okoh and Stein Haugen. A study of maintenance-related major accident cases in the 21st century. *Process Safety and Environmental Protection*, 92(4):346–356, July 2014.

[9] Marko Gerbec, Nora Balfe, Maria Chiara Leva, Steve Prast, and Micaela Demichela. Design of procedures for rare, new or complex processes: Part 1 –

An iterative risk-based approach and case study. *Safety Science*, 100:195–202, 2017.

[10] Marko Gerbec, Gabriele Baldissone, and Micaela Demichela. Design of procedures for rare, new or complex processes: Part 2 – Comparative risk assessment and CEA of the case study. *Safety Science*, 100:203–215, December 2017.

[11] Statista Industry Insight. Industrial Robotics - Worldwide | Statista Market Forecast, 2023.

[12] Amit Shukla and Hamad Karki. Application of robotics in onshore oil and gas industry—A review Part I. *Robotics and Autonomous Systems*, 75:490–507, January 2016.

[13] G.A. Raine and M.C. Lugg. ROV inspection of welds - A reality. *Insight: Non-Destructive Testing and Condition Monitoring*, 38(5):346–350, 1996.

[14] A. Hameed, A. Ordys, J. Możaryn, and A. Sibilska-Mroziewicz. Control System Design and Methods for Collaborative Robots: Review. *Applied Sciences (Switzerland)*, 13(1), 2023.

[15] Guilherme Deola Borges, Angélica Muffato Reis, Rafael Ariente Neto, Diego Luiz de Mattos, André Cardoso, Hatice Gonçalves, Eugenio Merino, Ana Colim, Paula Carneiro, and Pedro Arezes. Decision-Making Framework for Implementing Safer Human-Robot Collaboration Workstations: System Dynamics Modeling. *Safety*, 7(4):75, December 2021. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[16] Ana Pereira, José Dinis-Carvalho, Anabela Alves, and Pedro Arezes. How Industry 4.0 can enhance Lean practices. *FME Transactions*, 47(4):810–822, 2019.

[17] Luca Gualtieri, Erwin Rauch, and Renato Vidoni. Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, 67:101998, February 2021.

[18] S.J. Rasmussen and C.H. Houpis. Development, implementation and flight test of a MIMO digital flight control system for an unmanned research vehicle designed using quantitative feedback theory. *International Journal of Robust and Nonlinear Control*, 7(6):629–642, 1997.

[19] A D Swain and H E Guttmann. Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report. Technical Report NUREG/CR-1278, SAND-80-0200, 5752058, August 1983.

[20] Erik Hollnagel. *Cognitive reliability and error analysis method: CREAM*. Elsevier, Oxford ; New York, 1st ed edition, 1998.

[21] Oliver Sträter. Evaluation of human reliability on the basis of operational experience. *Gesellschaft für Anlagen und Reaktorsicherheit (GRS) mbH*, August 2000.

[22] Y.H.J. Chang and A. Mosleh. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 4: IDAC causal model of operator problem-solving response. *Reliability Engineering & System Safety*, 92(8):1061–1075, August 2007.

[23] Andreas Bye. Future needs of human reliability analysis: The interaction between new technology, crew roles and performance. *Safety Science*, 158:105962, February 2023.

[24] Yuan Sun, Idris Jeelani, and Masoud Gheisari. Safe human-robot collaboration in construction: A conceptual perspective. *Journal of Safety Research*, 86:39–51, September 2023.

[25] Antonio Giallanza, Giada La Scalia, Rosa Micale, and Concetta Manuela La Fata. Occupational health and safety issues in human-robot collaboration: State of the art and open challenges. *Safety Science*, 169:106313, January 2024.

[26] Shuo Yang, Micaela Demichela, Jie Geng, Ling Wang, and Zhangwei Ling. Risk-based performance assessment from fully manual to human-robot teaming in pressurized tank inspection operations. *Safety Science*, 176:106543, August 2024.

[27] Yang Shuo, Demichela Micaela, Geng Jie, and Tao Hongjie. Analysis of human and organization factors related accident reports based on natural language processing. *Chemical Engineering Transactions*, 90:745–750, 2022.

[28] Yang Shuo, Demichela Micaela, Geng Jie, Wang Ling, and Ling Zhangwei. Contributions and Consequences Coming from Human and Organizational Factors to the Accidents. *Chemical Engineering Transactions*, 91:523–528, 2022.

[29] Shuo Yang, Micaela Demichela, Jie Geng, Ling Wang, and Zhangwei Ling. A Data-Driven Bayesian Network of Management and Organizational Factors for Human Reliability Analysis in the Process Industry, April 2024.

[30] H. Toutenburg. Fisher, R. A., and F. Yates: Statistical Tables for Biological, Agricultural and Medical Research. 6th Ed. Oliver & Boyd, Edinburgh and London 1963. X, 146 P. Preis 42 s net. *Biometrische Zeitschrift*, 13(4):285–285, 1971. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bimj.19710130413.

[31] Kathleen B. Watson. Categorical data analysis. In Alex C. Michalos, editor, *Encyclopedia of quality of life and well-being research*, pages 601–604. Springer Netherlands, Dordrecht, 2014.

[32] Zhongdang Pan and Gerald Kosicki. Framing analysis: An approach to news discourse. *Political Communication*, 10(1):55–75, 1993.

[33] Steven C. Mallam, Aud Marit Wahl, and Jonas Aas. Reintroducing the sharp-end operator to organizational learning: How accident reports are used by maritime officers. *Safety Science*, 147:105632, March 2022.

[34] N. Tamim, S. Scott, W. Zhu, Y. Koirala, and M.S. Mannan. Roles of contractors in process safety. *Journal of Loss Prevention in the Process Industries*, 48:358–366, 2017.

[35] Esmaeil Zarei, Faisal Khan, and Mohammad Yazdi. A dynamic risk model to analyze hydrogen infrastructure. *International Journal of Hydrogen Energy*, 46(5):4626–4643, January 2021.

[36] Matthew Honnibal and Ines Montani. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. 2017.

[37] Johannes I. Single, Jürgen Schmidt, and Jens Denecke. Knowledge acquisition from chemical accident databases using an ontology-based method and natural language processing. *Safety Science*, 129:104747, September 2020.

[38] J. Reason. Human error: models and management. *Western Journal of Medicine*, 172(6):393–396, June 2000.

[39] Scott Shappell and Douglas Wiegmann. The Human Factors Analysis and Classification System-HFACS. January 2000.

[40] Ke-Lin Du and M. N. S. Swamy. *Neural Networks and Statistical Learning*. Springer London, London, 2019.

[41] S. Colombo, S. Nazir, A. Gallace, and D. Manca. Experiment-based decision making in complex systems. *Chemical Engineering Transactions*, 36:85–90, 2014.

[42] Seyedmohsen Hosseini and Dmitry Ivanov. Bayesian networks for supply chain risk, resilience and ripple effect analysis: A literature review. *Expert Systems with Applications*, 161:113649, December 2020.

[43] L. Mkrtchyan, L. Podofillini, and V. N. Dang. Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability Engineering & System Safety*, 139:1–16, July 2015.

[44] Nsimah J. Ekanem, Ali Mosleh, and Song-Hua Shen. Phoenix – A model-based Human Reliability Analysis methodology: Qualitative Analysis Procedure. *Reliability Engineering & System Safety*, 145:301–315, January 2016.

[45] Jing Xing, James Chang, and Jonathan DeJesus. NUREG-2198, "The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G)". page 372, 2021.

[46] Yuandan Li and Ali Mosleh. Dynamic simulation of knowledge based reasoning of nuclear power plant operator in accident conditions: Modeling and simulation foundations. *Safety Science*, 119:315–329, November 2019.

[47] Nancy Leveson. A new accident model for engineering safer systems. *Safety Science*, 42(4):237–270, April 2004.

[48] Erik Hollnagel. *FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-technical Systems*. CRC Press, November 2017. Google-Books-ID: je5HDwAAQBAJ.

[49] I.A. Herrera and R. Woltjer. Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis. *Reliability Engineering & System Safety*, 95(12):1269–1275, December 2010.

[50] Paulo Victor Rodrigues De Carvalho. The use of Functional Resonance Analysis Method (FRAM) in a mid-air collision to understand some characteristics of the air traffic management system resilience. *Reliability Engineering & System Safety*, 96(11):1482–1498, November 2011.

[51] Jue Li and Hongwei Wang. Modeling and analyzing multiteam coordination task safety risks in socio-technical systems based on FRAM and multiplex network: Application in the construction industry. *Reliability Engineering & System Safety*, 229:108836, January 2023.

[52] Yunlong Guo, Yongxing Jin, Shenping Hu, Zaili Yang, Yongtao Xi, and Bing Han. Risk evolution analysis of ship pilotage operation by an integrated model of FRAM and DBN. *Reliability Engineering & System Safety*, 229:108850, January 2023.

[53] Erik Hollnagel and Yushi Fujita. THE FUKUSHIMA DISASTER – SYSTEMIC FAILURES AS THE LACK OF RESILIENCE. *Nuclear Engineering and Technology*, 45(1):13–20, February 2013.

[54] Fabien Belmonte, Walter Schön, Laurent Heurley, and Robert Capel. Interdisciplinary safety analysis of complex socio-technological systems based on the functional resonance accident model: An application to railway trafficsupervision. *Reliability Engineering & System Safety*, 96(2):237–249, February 2011.

[55] Gui Fu, Bai Lu, and Xiuzhen Chen. Behavior based model for organizational safety management. *China Safety Science Journal*, 15(9):21–27, 2005.

[56] Micaela Demichela and Norberto Piccinini. Integrated Dynamic Decision Analysis (IDDA): an Advanced Tool for Risk Analysis. In Cornelia Spitzer, Ulrich Schmocker, and Vinh N. Dang, editors, *Probabilistic Safety Assessment and Management*, pages 2956–2961. Springer London, London, 2004.

[57] S. Clementel and R. Galvagni. The use of the event tree in the design of nuclear power plants. *Environment International*, 10(5):377–382, January 1984.

[58] Micaela Demichela and Gianfranco Camuncoli. Risk based decision making. Discussion on two methodological milestones. *Journal of Loss Prevention in the Process Industries*, 28:101–108, April 2014.

[59] Gabriele Baldissone, Davide Fissore, and Micaela Demichela. Catalytic after-treatment of lean VOC–air streams: Process intensification vs. plant reliability. *Process Safety and Environmental Protection*, 100:208–219, March 2016.

[60] Jon Norberg and Graeme Cumming. *Complexity theory for a sustainable future*. Columbia University Press, 2008.

[61] Antonio Carlos Braz and Adriana Marotti de Mello. Circular economy supply network management: A complex adaptive system. *International Journal of Production Economics*, 243:108317, January 2022.

[62] Kristiaan P. W. Kok, Anne M. C. Loeber, and John Grin. Politics of complexity: Conceptualizing agency, power and powering in the transitional dynamics of complex adaptive systems. *Research Policy*, 50(3):104183, April 2021.

[63] Melanie Mitchell. *Complexity: A guided tour*. Oxford university press, 2009.

[64] Fan Ouyang, Weiqi Xu, and Mutlu Cukurova. An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning*, 18(1):39–66, March 2023.

[65] John Van Aerde, Marcio M Gomes, Meredith Giuliani, Brent Thoma, and Susan Lieff. Complex adaptive systems in CanMEDS 2025. *Canadian Medical Education Journal*, 14(1):50–53, March 2023.

[66] John H. Holland. *Hidden order: how adaptation builds complexity*. Helix books. Addison-Wesley, Reading, Mass, 1995.

[67] W. Brian Arthur. Complexity and the Economy. *Science*, 284(5411):107–109, April 1999. Publisher: American Association for the Advancement of Science.

[68] Vito Albino, Nunzia Carbonara, and Ilaria Giannoccaro. Innovation in industrial districts: An agent-based simulation model. *International Journal of Production Economics*, 104(1):30–45, November 2006.

[69] Thomas Y Choi, Kevin J Dooley, and Manus Rungtusanatham. Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, 19(3):351–366, May 2001.

[70] Li An, Volker Grimm, Abigail Sullivan, B. L. Turner II, Nicolas Malleson, Alison Heppenstall, Christian Vincenot, Derek Robinson, Xinyue Ye, Jianguo Liu, Emilie Lindkvist, and Wenwu Tang. Challenges, tasks, and opportunities in modeling agent-based complex systems. *Ecological Modelling*, 457:109685, October 2021.

[71] Malvina Roci, Niloufar Salehi, Saman Amir, Sayyed Shoaib-ul Hasan, Farazee M. A. Asif, Aleš Mihelič, and Amir Rashid. Towards circular manufacturing systems implementation: A complex adaptive systems perspective using modelling and simulation as a quantitative analysis tool. *Sustainable Production and Consumption*, 31:97–112, May 2022.

[72] Nicolae Lobontiu. Chapter 10 - Coupled-Field Systems. In Nicolae Lobontiu, editor, *System Dynamics for Engineering Students (Second Edition)*, pages 487–540. Academic Press, Boston, January 2018.

[73] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl_3):7280–7287, May 2002. Publisher: Proceedings of the National Academy of Sciences.

[74] Michael W. Macy and Robert Willer. From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, 28(Volume 28, 2002):143–166, August 2002. Publisher: Annual Reviews.

[75] Sherif Abdelkhalek and Tarek Zayed. Simulation-based planning of concrete bridge deck inspection with non-destructive technologies. *Automation in Construction*, 119:103337, November 2020.

[76] Stef Janssen, Alexei Sharpanskykh, and Richard Curran. AbSRiM: An Agent-Based Security Risk Management Approach for Airport Operations. *Risk Analysis*, 39(7):1582–1596, July 2019.

[77] Xing Guo, Qiang Feng, Dongming Fan, Zili Wang, Yi Ren, Bo Sun, and Dezhen Yang. An agent-based dynamic reliability modeling method for multistate systems considering fault propagation: A case study on subsea Christmas trees. *Process Safety and Environmental Protection*, 158:20–33, February 2022.

[78] Yali Wu, Gui Fu, Meng Han, Qingsong Jia, Qian Lyu, Yuxin Wang, and Zhirong Wu. Comparison of the theoretical elements and application characteristics of STAMP, FRAM, and 24Model: A major hazardous chemical explosion accident. *Journal of Loss Prevention in the Process Industries*, 80:104880, December 2022.

[79] B. Kirwan and L.K. Ainsworth. *A guide to task analysis: The task analysis working group*. A Guide To Task Analysis: The Task Analysis Working Group. 1992. Pages: 413.

[80] Helena Anna Frijns, Oliver Schürer, and Sabine Theresia Koeszegi. Communication Models in Human–Robot Interaction: An Asymmetric MODel of ALterity in Human–Robot Interaction (AMODAL-HRI). *International Journal of Social Robotics*, 15(3):473–500, March 2023.

[81] Lin-Xiu Hou, Ran Liu, Hu-Chen Liu, and Shan Jiang. Two decades on human reliability analysis: A bibliometric analysis and literature review. *Annals of Nuclear Energy*, 151:107969, February 2021.

[82] Jie Geng, Salvina Mure, Gabriele Baldissone, Gianfranco Camuncoli, and Micaela Demichela. Human error probability estimation in atex-hmi area classification: from therp to fuzzy cream. *Chemical Engineering Transactions*, 43:1243–1248, 2015.

[83] Tuqiang Zhou, Chaozhong Wu, Junyi Zhang, and Di Zhang. Incorporating CREAM and MCS into fault tree analysis of LNG carrier spill accidents. *Safety Science*, 96:183–191, July 2017.

[84] Ross Owen Phillips and Fridulv Sagberg. What did you expect? CREAM analysis of hazardous incidents occurring on approach to rail signals. *Safety Science*, 66:92–100, July 2014.

[85] Xi Zheng, Matthew L. Bolton, Christopher Daly, and Elliot Biltekoff. The development of a next-generation human reliability analysis: Systems analysis for formal pharmaceutical human reliability (SAFPH). *Reliability Engineering & System Safety*, 202:106927, October 2020.

[86] Glenn Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.

[87] Arthur P. Dempster. Upper and Lower Probabilities Induced by a Multivalued Mapping. In Roland R. Yager and Liping Liu, editors, *Classic Works of the Dempster-Shafer Theory of Belief Functions*, volume 219, pages 57–72. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. Series Title: Studies in Fuzziness and Soft Computing.

[88] Ramisetty Kavya, Christopher Jabez, and Panda Subhrakanta. A new belief interval-based total uncertainty measure for Dempster-Shafer theory. *Information Sciences*, 642:119150, September 2023.

[89] Sukru Ilke Sezer, Emre Akyuz, and Ozcan Arslan. An extended HEART Dempster–Shafer evidence theory approach to assess human reliability for the gas freeing process on chemical tankers. *Reliability Engineering & System Safety*, 220:108275, April 2022.

[90] Tasneem Bani-Mustafa, Zhiguo Zeng, Enrico Zio, and Dominique Vasseur. A new framework for multi-hazards risk aggregation. *Safety Science*, 121:283–302, January 2020.

[91] Runze Liu, Liguo Fei, and Jianing Mi. An evidential MULTIMOORA approach to assessing disaster risk reduction education strategies under a heterogeneous linguistic environment. *International Journal of Disaster Risk Reduction*, 78:103114, August 2022.

[92] Z.S. Givi, M.Y. Jaber, and W.P. Neumann. Modelling worker reliability with learning and fatigue. *Applied Mathematical Modelling*, 39(17):5186–5199, September 2015.

[93] Bruno Siciliano, Lorenzo Sciavicco, Luigi Villani, and Giuseppe Oriolo. *Robotics*. Advanced Textbooks in Control and Signal Processing. Springer London, London, 2009.

[94] J. Carlson, R.R. Murphy, and A. Nelson. Follow-up analysis of mobile robot failures. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, pages 4987–4994 Vol.5, New Orleans, LA, USA, 2004. IEEE.

[95] Guang Zou and Athanasios Kolios. Quantifying the value of negative inspection outcomes in fatigue maintenance planning: Cost reduction, risk mitigation and reliability growth. *Reliability Engineering & System Safety*, 226:108668, October 2022.

[96] M. Karthick, I. Ahmed Irfan, K. Mohan Babu, and T. Satyanarayana. Design and fabrication of multifunctional remote control vehicle. *International Journal of Mechanical and Production Engineering Research and Development*, 2018(Special Issue):164–169, 2018.

[97] Z. Li, D. Wang, H. Gao, H. Cao, Y. Zhao, Z. Miao, Z. Yang, and W. He. Application of double click to prepare near-infrared absorbing dye for photothermal tuning of cholesteric liquid crystal. *Reactive and Functional Polymers*, 151, 2020.

[98] Dorota Stadnicka, Paweł Litwin, and Dario Antonelli. Human factor in industry of the future: Knowledge acquisition and motivation. *FME Transactions*, 47(4):823–830, 2019.

[99] Qiang Wu, Xunpen Qin, Kang Dong, Aixian Shi, and Zeqi Hu. A learning-based crack defect detection and 3D localization framework for automated fluorescent magnetic particle inspection. *Expert Systems with Applications*, 214:118966, March 2023.

[100] Yonggui Chen, Ping Tang, Shuai Kong, Sunting Yan, Chenfeng Guan, and Wei Tian. Automation Solution to Wall-climbing Robot Working for Large-scale Metal Spherical Tanks. *Highlights in Science, Engineering and Technology*, 9:168–173, September 2022.

[101] David L. Olson and Dursun Delen. Performance Evaluation for Predictive Modeling. In David L. Olson and Dursun Delen, editors, *Advanced Data Mining Techniques*, pages 137–147. Springer, Berlin, Heidelberg, 2008.

[102] William T Fine. Mathematical Evaluations for Controlling Hazards. Technical Report AD0722011, NAVAL ORDNANCE LAB WHITE OAK MD, January 1970. Section: Technical Reports.

[103] Joseph Jonathan Magoua, Fei Wang, Nan Li, and Dongping Fang. Incorporating the human factor in modeling the operational resilience of interdependent infrastructure systems. *Automation in Construction*, 149:104789, May 2023.

[104] Minghui Wu, Jia-Rui Lin, and Xin-Hao Zhang. How human-robot collaboration impacts construction productivity: An agent-based multi-fidelity modeling approach. *Advanced Engineering Informatics*, 52:101589, April 2022.

[105] S. Ehsan Hashemi-Petroodi, Simon Thevenin, Sergey Kovalev, and Alexandre Dolgui. Operations management issues in design and control of hybrid human-robot collaborative manufacturing systems: a survey. *Annual Reviews in Control*, 49:264–276, January 2020.

[106] Michael A. Goodrich and Alan C. Schultz. Human–Robot Interaction: A Survey. *Foundations and Trends® in Human–Computer Interaction*, 1(3):203–275, January 2008. Publisher: Now Publishers, Inc.

[107] S. S. Shapiro and M. B. Wilk. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591, December 1965.

[108] H. B. Mann and D. R. Whitney. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1):50–60, March 1947.

[109] J. D. Lee and K. A. See. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1):50–80, January 2004.

[110] Ning Wang, David V. Pynadath, K. V. Unnikrishnan, Santosh Shankar, and Chirag Merchant. Intelligent Agents for Virtual Simulation of Human-Robot Interaction. In Randall Shumaker and Stephanie Lackey, editors, *Virtual, Augmented and Mixed Reality*, Lecture Notes in Computer Science, pages 228–239, Cham, 2015. Springer International Publishing.

[111] Erkan Dogan, Mehmet Ali Yurdusev, Sadik Alper Yildizel, and Gokhan Calis. Investigation of scaffolding accident in a construction site: A case study analysis. *Engineering Failure Analysis*, 120:105108, February 2021.

[112] Xin Li, Yong Guo, Fan-liang Ge, and Fu-qiang Yang. Human reliability assessment on building construction work at height: The case of scaffolding work. *Safety Science*, 159:106021, March 2023.

[113] G. Klutke, P.C. Kiessler, and M.A. Wortman. A critical look at the bathtub curve. *IEEE Transactions on Reliability*, 52(1):125–129, March 2003.

[114] Yun Yang, Yao Yang, Long Li, Cuili Chen, and Zhou Min. Automatic Defect Identification Method for Magnetic Particle Inspection of Bearing Rings Based on Visual Characteristics and High-Level Features. *Applied Sciences*, 12(3):1293, January 2022. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[115] N. J. Shipway, P. Huthwaite, M. J. S. Lowe, and T. J. Barden. Using ResNets to perform automated defect detection for Fluorescent Penetrant Inspection. *NDT & E International*, 119:102400, April 2021.

[116] N.J. Shipway, T.J. Barden, P. Huthwaite, and M.J.S. Lowe. Automated defect detection for Fluorescent Penetrant Inspection using Random Forest. *NDT & E International*, 101:113–123, January 2019.

[117] Julien Cambon, Franck Guarnieri, Jop Groeneweg, E. Scholten, J. Hinrichs, and G.E. Lancioni. Bringing Tripod Delta to France for the analysis of organizational factors. In Edited by Carlos Guedes Soares and Enrico Zio, editors, *European Safety and Reliability Conference (ESREL 2006)*, volume 1-3 of *Proceedings and Monographs in Engineering, Water and Earth Sciences*, pages 267–275, Estoril, Portugal, 2006. Taylor & Francis.

[118] A.r. Hale. Safety management in production. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 13(3):185–201, 2003. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hfm.10040.

[119] Great Britain and Health and Safety Executive. *Developing process safety indicators: a step-by-step guide for chemical and major hazard industries*. HSE Books, Sudbury, Suffolk, 2006. OCLC: 1017839900.

[120] K. Øien. A framework for the establishment of organizational risk indicators. *Reliability Engineering & System Safety*, 74(2):147–167, November 2001.

[121] Katrina M Groth. A DATA-INFORMED MODEL OF PERFORMANCE SHAPING FACTORS FOR USE IN HUMAN RELIABILITY ANALYSIS. page 243, 2009.

[122] Trond Kongsvik, Petter Almklov, and Jørn Fenstad. Organisational safety indicators: Some conceptual considerations and a supplementary qualitative approach. *Safety Science*, 48(10):1402–1411, December 2010.

[123] M. Paradies, L. Unger, P. Haas, and M. Terranova. Development of the NRC‘s Human Performance Investigation Process (HPIP). Volume 3, Development documentation. Technical Report NUREG/CR–5455-Vol.3, SI–92-101-Vol.3, 10106211, October 1993.

Appendix A

Organizational factors

Table A.1 Definition of Organizational factors in the literature

Factors in eMARS	Synonyms	Description	Literature source
Design of plant / equipment / system	Design	Ergonomically poor design of tools or equipment	(Cambon et al., 2006)[117]
	Design	Technical design of plant and hardware and its safe modification to provide optimal safety	(Hale, 2003)[118]
	Plant design	When selecting suitable equipment, consider the following: standards and codes, compatibility of materials with products, anticipated duty and degradation methods, and pressure systems. life expectancy, electrical integrity and equipment bonding, and ease of inspection and maintenance.	(Great Britain & Health and Safety Executive,2006)[119]

	Design	refers to the physical construction and assembly of the process- and other equipment. The design must be such that the installation can be operated and maintained without causing leaks.	(Øien, 2001)[120]
Maintenance / Repair	Maintenance	no or inadequate performance of maintenance tasks and repairs, bad planning	(Cambon et al., 2006)[117]
	Inspection and maintenance	The specification of scope and frequency of the inspection and maintenance system. This should be based on how safety-critical the item is and on the degree of the challenge presented to the system integrity or compliance with the manufacturer's or supplier's instructions.Safety-critical plant and equipment (i.e. flexi hoses, couplings, pump valves, flanges, fixed pipes, bulk tanks) are inspected for wear and damage or malfunction within the specified period.Faults are fixed within specified timescales, and repairs and improvements meet plant design standards.A log of findings is kept – enabling trending.	(Great Britain & Health and Safety Executive, 2006)[119]
Management attitude problem	Organization Climate	Refers to the working atmosphere within the organization (e.g., structure, policies, culture)	(Shappell and Wiegmann, 2000)[39]

	Top management culture	system to manage conflicts between safety and other company goals explicitly, for example, in production and maintenance planning, purchasing, design, and so forth	(Hale, 2003)[118]
	Safety culture	characterizes the organizational attitude, values, and beliefs toward workers and public safety	(Groth, 2009)[121]
Management organization inadequate	Organization inadequate	shortcomings in the organizational structure, organization philosophy, and management strategies	(Cambon et al., 2006)[117]
	Organization inadequate	The quality of the roles and responsibilities of team members, additional support, communication systems, Safety Management System, instructions and guidelines for externally oriented activities, the role of external agencies, etc.	(Hollnagel, 1998)[20]
Organized procedures	procedures	insufficient quality or availability of procedures, manuals, and written instructions	(Cambon et al., 2006)[117]
	procedures, goals, plans, and rules	specifying what to achieve in safety and/or how to achieve it	(Hale, 2003)[118]
	Operation procedures	Procedures contain correct scope (key actions and tasks including emergency action) and/or sufficient detail.Procedures are clearly written/easily understood.Procedures are kept up to date.	(Great Britain and Health and Safety Executive, 2006)[119]

Procedures, JSA, guidelines, and instructions	refer to all written and oral information describing how to perform the operation and maintenance tasks correctly and safely. The main emphasis is on the task information necessary to avoid leaks.	(Øien, 2001)[120]
Procedures and document	the written remedies and illustrations that describe operational/maintenance routines and plant/installation design/status	(Kongsvik et al., 2010)[122]
Procedures	Procedures are explicit, step-by-step instructions for performing a task.	(Groth, 2009)[121]
Procedures, guidelines, and instructions	This PIF refers to the availability and usefulness of operating procedures, guidance, and instructions (including protocols). Procedures, guidance, and instructions (PGIs) should be validated for their applicability and usefulness. Following PGIs should lead to the success of important human actions.	(Xing et al., 2021)[45]
Formal Written guidance	including maintenance manuals, surveillance procedures, operating procedures, and emergency operating procedures are provided to workers or supervisors.	(Paradies et al., 1993)[123]
Availability of procedures /plans	Procedures and plans include operating and emergency procedures, familiar patterns of response heuristics, routines, etc. Procedures and plans include operating and emergency procedures, familiar patterns of response heuristics, routines, etc.	(Hollnagel, 1998)[20]

	Availability of procedures	Refer to the availability and quality of the explicit step-by-step instructions the crew needs to perform a task. Ideally, the crew should commit no errors when they are following the procedure correctly. However, procedures could be written incorrectly, leading the crew to make errors even with the right intent. This group is made up of two level 2 PIFs, namely: Procedure Quality and Procedure Availability.	(Ekanem et al., 2016)[44]
Staffing	Staff competence	Information and training covering:hazardous properties of products;ship-to-shore communication systems;pre-transfer checks;product transfer controls and monitoring; post-transfer checks;emergency actions. Job-specific knowledge and relevant experience of:substances; work processes; hazards; and emergency actions.	(Great Britain & Health and Safety Executive, 2006)[119]
	Personnel	refers to the way that the organization hires and assigns tasks to personnel	(Groth, 2009)[121]
	Staff	Refers to having adequate, qualified personnel to perform the required tasks. Staffing includes the number of personnel, their skill sets, job qualifications, and staffing structure (individual and team roles and responsibilities). Adequate and qualified staff is normally expected.	(Xing et al., 2021)[45]

	Manning	refer to how many and what kinds of people perform which types of jobs.	(Swain & Guttmann, 1983)[19]
	Resource Management	Refers to the organizational-level decision-making regarding allocating and maintaining organizational assets (e.g., human resources, monetary/budget resources, equipment/facility recourse). Organizational Climate: Refers to the working atmosphere within the organization (e.g., structure, policies, culture). Operational Process: Refers to organizational decisions and rules that govern the everyday activities within an organization (e.g., operations, procedures, oversight).	(Shappell & Wiegmann, 2000)[39]
Training / instruction	Training	inadequate planning, the ineffectiveness of training, insufficient competence or experience of personnel	(Cambon et al., 2006)[117]
	Training	refers to the knowledge and experience imparted to the personnel by the utility. Training includes training courses' content, the training courses' scheduling, and the training frequency.	(Groth, 2009) [121]

Training	refers to training that personnel receive to perform their tasks. This consideration includes personnel's work-related experience, whether they have been trained on the type of the event, and the amount of time passed since training and training on the specific systems involved in the event. It is expected that adequate training is required for professional staff.	(Xing et al., 2021)[45]
Training	review training documentation and training records to assess the adequacy of the training program for the tasks related to the event.Interview training department personnel. Discuss specific problems with worker/supervisor knowledge and skills identified during interviews with workers and supervisors.	(Paradies et al., 1993)[123]
Adequacy of training and experience	The level and quality of training provided to operators as familiarization with new technology, refreshing old skills, etc. It also refers to the level of operational experience.	(Hollnagel, 1998)[20]
Training/competence	refers to the training and competence that is necessary for the operating personnel to carry out their jobs without causing any leaks. This covers both general system knowledge and specific skills required for operational and maintenance tasks	(Øien, 2001)[120]

Direct Supervision	Direct supervision	Direct supervision serves as the link between management and the team members. The direct supervisor can be seen as a member of the team, albeit a member with additional authority and responsibility.	(Groth, 2009)[121]
	supervision	First or second-line managers of production/maintenance works, including preparation and supervision during work.	(Paradies et al., 1993)[123]
User-unfriendliness (apparatus, system, etc.)	Ergonomic	A user-friendly and ergonomically responsible interface in all life-cycle phases	(Hale, 2003)[118]
	Human system interface	The Human-System Interface PSF covers how information is communicated between humans and machines.	(Groth, 2009)[121]
	Human machine interface	refers to indications (e.g., displays, indicators, labels) and controls used by personnel to execute actions on systems	(Xing et al., 2021)[45]
	Human system interface	Refers to the ways and means of interaction between the crew and the system. This PIF covers the quality (usability, ergonomics, physical access, etc.) of the HSI regarding system output and the crew's input.	(Ekanem et al., 2016)[44]

Adequacy of MMI and operational support	The Man-Machine Interface generally includes the information available on control panels, computerized workstations, and operational support provided by specifically designed decision aids.	(Hollnagel, 1998)[20]
Human engineering	Criteria that support reliable human performance and will result in people doing tasks in a consistently correct manner.Including human-machine interface, labels less than adequate, arrangement/placement, instrument/displays less than adequate, controls less than adequate, monitoring alertness less than adequate, and unit differences.	(Paradies et al., 1993)[123]

Appendix B

Data for IDDA

B.1 Basic information of case study teams

Team 1: The Chemical Plant Tank Maintenance Team owns the spherical tank and will inspect and provide the working conditions. 1. They usually have an electricity permit management system, fixed organization structure, and staff. 2. They most of the time need to work in the field of chemical storage equipment area. 3. They usually provide good state sufficient equipment and tools to workers. 4. They always have printed procedures and temporary project plans with schedule arrangements. 5. Usually, workers in the team need to do different tasks simultaneously. 6. The workers in this team usually have adequate time to perform the task. 7. They usually work during the daytime. 8. They always receive regular training in the plant 9. They usually have good communication within the team and acceptable communication with other teams.

Team 2: Scaffold Building Team, which needs to build the scaffold for the others to access the spherical tank area. 1. They usually temporarily team and staff, haired and supervised by a chemical plant. 2. Most of the time, they work inside the spherical tank as a confined space. 3. They usually provide good state sufficient equipment and tools to workers. 4. They sometimes have printed procedures but no detailed project plans. 5. Sometimes, workers in the team need to do different tasks simultaneously. 6. The workers in this team sometimes have adequate time to perform the task. 7. They usually work during the daytime. 8. They always receive temporary training from the plant and have the required certificate for working. 9. They usually have good communication within the team and acceptable communication with other teams.

Team 3: Polish Workers Team, which needs to polish the weld joints of the tank. 1. They usually temporarily team and staff, haired and supervised by a chemical plant. 2. they usually work inside the spherical tank in a confined space under high noise. 3. They usually provide good state sufficient equipment and tools to workers. 4. They usually have no procedures or detailed project plans but have the required polish quality criteria. 5. Sometimes, workers in the team need to do different tasks simultaneously. 6. The workers in this team sometimes have adequate time to perform the task. 7. They usually work during the daytime. 8. They always receive temporary training from the plant and have the required certificate for working. 9. They usually have good communication within the team and sometimes less acceptable communication with other teams.

Team 4: Full Manual Technician Team, which must perform the fully manual magnetic particle inspection inside the tank. 1. They belong to a government institute; the team has a temporary organizational structure and staff. They deliver the requirements to front-line technicians efficiently and orally. 2. most of the time, work inside the spherical tank as a confined space, 3. They always provide good state sufficient equipment and tools to workers. 4. They usually have printed procedures and project plans with schedule arrangement, but not so detailed. 5. Always, workers in the team do not need to do different tasks simultaneously. 6. The workers in this team usually have adequate time to perform the task. 7. They always work during the daytime. 8. They always receive regular training at the institute, temporary training from the plant, and have the required certificate for inspection work. 9. They usually have good communication within the team and good communication with other teams.

Team 5: Human-Robot Collaborative Technician Team, which needs to perform both the polish and magnetic work with robot and human monitoring. 1. They belong to a government institute; the team has a temporary organizational structure and staff. They deliver the requirements to front-line technicians efficiently and orally. 2. They sometimes need to work in the field of chemical storage equipment area for robot settings. Mostly, they work in the control room. 3. They always provide good state sufficient equipment and tools to workers. But the robots are in the early use stage and, therefore, may have some problems that need to be fixed. 4. They usually have some printed procedures and project plans with schedule arrangements, but not so detailed. However, some robot-related procedures are not well-detailed and developed. 5. Occasionally, workers in the team need to do different tasks at the same time. 6. The workers in this team usually have adequate time to perform the task. 7. They always work during the daytime. 8. They always receive regular training at the institute and temporary training from the plant to have the required certificate for inspection work. However, their experience with robot settings and control is acceptable but not high. 9. They usually have good communication within the team and good communication with other teams.

B.2 Task analysis

Table B.1 A list of subtasks and dependencies for the FM inspection process

NO.	Subtasks	Actor	Location	Failure mode	Input	Output	D	W	O
1.1	Make the inspection plan	Technician supervisor1	Office room	P2		Plan			
1.2	Check the inspection plan	Technician supervisor2	Office room	O2,I1	1.1	Plan	*		
1.3	Insert the blind disc	Plant Operator1	Outside tank	E1		Isolated state			
1.4	Check the isolation	Plant Operator2	Outside tank	O2,I1	1.3	Isolated state	*		
1.5	Electricity environment setting	Plant Operator1	Manhole	E3		Electricity environment			*
1.6	Start and keep venting for 3 days	Plant Operator1	Manhole	E2		Ventilation finished			*
1.7	Take the air sample	Plant Operator1	Manhole	E1		Air sample			
1.8	Test the air concentration	Plant lab	Plant lab	E1	1.7	air component concentrate	*		
1.9	Perform the safety training	Plant safety manager	Outside tank	E5		Trained operators			
2.1	Check all materials and tools availability	Construction supervisor1	Outside tank	E5		Qualified tools			
2.2	Wear PPE	Construction team1	Outside tank	E5		Protected state			
2.3	Test the air concentration	Construction supervisor1	Manhole	E1	2.1	Air component concentrate			
2.4	Enter the tank with a gas detector	Construction team1	Outside tank	E5	2.2,2.3		*		*
2.5	Climb up ladder	Construction team1	Manhole	E3	2.2		*		*
2.6	Build the scaffold	Construction team1	Manhole	E1	2.2	Scaffold	*		*
2.7	Handoff to next shift after two hours	Construction team1	Manhole	E2			*		*
2.8	Check scaffold quality	Construction supervisor1-2	Inside tank	O2,I1	2.6	Scaffold qualified	*		
2.9	Cross-check scaffold quality	Plant safety manager	Inside tank	O2,I1	2.8		*		
3.1	Check all materials and tools availability	Polish supervisor1	Outside tank	E5		Qualified tools			
3.2	Wear PPE	Polish team1	Outside tank	E5		Protected state			
3.3	Test the air concentration	Polish supervisor1	Manhole	E1	3.1	Air component concentrate			
3.4	Enter the tank with a gas detector	Polish team1	Outside tank	E5	3.2,3.3		*		*
3.5	Climb up ladder	Polish team1	Manhole	E3	2.9		*		*
3.6	Build the scaffold	Polish team1	Manhole	E1		Scaffold	*	*	*
3.7	Handoff to next shift after two hours	Polish team1	Manhole	E2			*	*	*
3.8	Check scaffold quality	Polish supervisor1-2	Inside tank	O2,I1	3.6	Polished joints qualified	*	*	*
3.9	Cross-check scaffold quality	Plant safety manager	Inside tank	O2,I1	3.8	Polished joints qualified	*	*	*
4.1	Check all materials and tools availability	Technician supervisor1	Outside tank	E5		Qualified tools	*		
4.2	Perform magnetic liquid sensitivity test	Technician team1	Outside tank	E5		Qualified material	*	*	
4.3	Wear PPE	Technician team1	Manhole	E5	4.1	Protected state			
4.4	Test the air concentration	Technician supervisor1	Manhole	E1	4.1	Air component concentrate			
4.5	Enter the tank with a gas detector	Technician team1	Inside tank	E5	4.4		*		*
4.6	Climb up ladder	Technician team1	Inside tank	E3	2.11		*		*
4.7	Spray the liquid toward the weld joint	Technician1	Inside tank	E3	3.11	Wet weld joints	*	*	*
4.8	Magnetic with X posture	Technician1	Inside tank	E1	4.7	Magnetic patterns	*	*	*
4.9	Irradiate with UV light and identify a crack	Technician1	Inside tank	O2,I2	4.8	Marked cracks	*	*	*
4.10	Mark the crack	Technician1	Inside tank	E3					
4.11	Handoff to next shift after two hours	Technician team1	Inside tank	E2			*	*	*
4.12	Check identified cracks	Technician supervisor1-2	Inside tank	O2,I1		Confirmed cracks	*	*	*
5.1	Check all materials and tools availability	Construction supervisor1	Outside tank	E5		Qualified tools			
5.2	Wear PPE	Construction team1	Outside tank	E1	5.1	Protected state			
5.3	Test the air concentration	Construction supervisor1	Manhole	E1	5.1	Air component concentrate			
5.4	Enter the tank with a gas detector	Construction team1	Outside tank	E5	5.3				
5.5	Climb up ladder	Construction team1	Manhole	E3			*		*
5.6	Build the scaffold	Construction team1	Manhole	E3			*		*
5.7	Handoff to next shift after two hours	Construction team1	Inside tank	E2			*	*	*
6.1	Detect the low O2 alarm	Operators	Manhole	O3,I1					*
6.2	Perform rescue procedures	Plant safety manager	Outside tank	P2,E2	1.4				*
6.3	Check inside members are safe every 15min	Plant safety manager	Outside tank	E5					

Table B.2 A list of subtasks and dependencies for the HRT inspection process

NO.	Subtasks	Actor	Location	Failure mode	Input	Output	D	W	O
1.1	Make the inspection plan	Technician supervisor1	Office room	P2		Plan			
1.2	Check the inspection plan	Technician supervisor2	Office room	O2,I1	1.1	Plan	*		
1.3	Insert the blind disc	Plant Operator1	Outside tank	E1		Isolated state			
1.4	Check the isolation	Plant Operator2	Outside tank	O2,I1	1.3	Isolated state	*		
1.5	Electricity environment setting	Plant Operator1	Manhole	E3		Electricity environment			*
1.6	Start and keep venting for 3 days	Plant Operator1	Manhole	E2		Ventilation finished			*
1.7	Take the air sample	Plant Operator1	Manhole	E1		Air sample			
1.8	Test the air concentration	Plant lab	Plant lab	E1	1.7	air component concentrate	*		
1.9	Perform the safety training	Plant safety manager	Outside tank	E5		Trained operators			
2.1	Check all materials and tools	Technician1-2	Outside tank	E5		Qualified tools	*		
2.2	Wear PPE	Technician1-4	Outside tank	E5	2.1	Protected state			
2.3	Test the air concentration	Technician1	Manhole	E1	2.1	Air component concentrate			
2.4	Enter the tank with a gas detector	Technician1-2	Manhole	E5	2.3				*
2.5	Connect the robot with electricity	Technician1-2	Manhole	E3		Robot with electricity			*
2.6	Climb the ladder and at the manhole	Technician3-4	Tank outside surface	E3					*
2.7	Connect the polish robot with a safety rope	Technician3-4	Manhole	E1		Robot protected			*
2.8	Set the polish robot parameters	Technician1-2	Outside tank	O2,E3		Robot set			*
2.9	Start and monitor the polish robot	Technician1-2	Office room	O2,I1		Robot started	*	*	
2.10	Handoff to next shift after two hours	Technician1-2	Office room	E2					
2.11	Polish robot moves along the weld joint	Polish robot	Inside tank	E1			*	*	
2.12	Polish robot polishes with proper force	Polish robot	Inside tank	E1		Polished joints	*	*	
2.13	Polish robot videos of the Polish process	Polish robot	Inside tank	O3		Video record	*	*	
2.14	Recovery from the wrong situation	Technician1-2	Office room	O2,I1,P2,E3	2.13,2.14,2.15		*		
2.15	Detect the iron pieces cumulated in time	Technician team	Manhole	O3	3.8	Robot cleaned	*		
3.1	Check all materials and tools	Technician3-4	Outside tank	E5		Qualified tools	*		
3.2	Wear PPE	Technician1-4	Outside tank	E5	3.1	Protected state			
3.3	Test the air concentration	Technician1-4	Manhole	E1	3.1	Air component concentrate			
3.4	Enter the tank with a gas detector	Technician1-2	Manhole	E5	3.3				*
3.5	Connect the robot with electricity	Technician1-2	Manhole	E3		Robot with electricity			*
3.6	Climb the ladder and at the manhole	Technician3-4	Tank outside surface	E3					*
3.7	Connect the magnetic robot with a safety rope	Technician3-4	Manhole	E1		Robot protected			*
3.8	Set the magnetic robot parameters	Technician3-4	Inside tank	O2,E3		Robot set	*		*
3.9	Start and monitor the magnetic robot	Technician1-2	Office room	O2,I1		Robot started	*	*	
3.10	Handoff to next shift after two hours	Technician1-2	Office room	E2					
3.11	Magnetic robot moves along the weld joint	Magnetic robot	Inside tank	E1			*	*	
3.12	Magnetic robot magnetics with proper speed	Magnetic robot	Inside tank	E1		Magnetic joints	*	*	
3.13	Magnetic robot videos of the magnetic process	Magnetic robot	Inside tank	O3		Video record	*	*	
3.14	Recovery from the wrong situation	Technician1-2	Office room	O2,I1,P2,E3	3.13,3.14,3.15	*			
3.15	Magnetic robot identifies a crack	Magnetic robot	Inside tank	O2,I2	2.14,3.14	Identified crack		*	
3.16	Magnetic robot sends an alarm	Magnetic robot	Inside tank	E5		Crack alarm		*	
3.17	Detect the alarm	Technician Team	Office room	O2,I1	3.17	Acknowledged alarm		*	
3.18	Confirm/deny the crack	Technician Team	Office room	O2,I1	3.16	Confirmed/denied crack		*	
4.1	Detect the low O2 alarm	Plant safety manager	Outside tank	O2		Acknowledged alarm			
4.2	Perform rescue procedure	Plant safety manager	Outside tank	P2,E2	1.4				*

B.3 Input file of IDDA for FM process

! Task 1 Prepare and Authorise the Work

101 1, 0, 1., 102 132,' The spherical tank inside is empty and purged ' ' Yes' 'No'

L 101 1, 132 1

102 1, 0.00192, 1., 103 110,' Insert the blind disc ' 'Yes' 'No'

110 1,0.025,1., 103 111, 'Recovery' ' Yes ' 'No '

L 110 1, 111, 1

111 1, 0, 1., 103 103,' Leak LPG into the spherical ' 'No' ' Yes '

A 111 1, 103 115 115

A 111 1, 105 120 120

L 111 1, 120 1

103 1, 0, 1., 104 104,' Open the manhole ' 'Yes' 'No'

115 1,0.00768,1., 116 116, 'Ignition source' 'No' 'Yes'

L 115 1, 116 1

116 1,0,1., 104 131, 'Flash Fire burnt the operator' 'No' 'Yes'

L 116 1, 131 1

104 1, 0.000032, 1., 105 131, 'Set up electricity environment get shocked ' ' No ' ' Yes '

L 104 1, 131 1

105 1, 0.000192, 1., 106 106,' Venting with venting machine for more than 3 days' 'Yes' 'No'

L 105 1, 106 1

120 1,0,1., 106 132 'LPG Release' 'No' 'Yes'

L 120 1, 132 1

106 1, 0, 1., 107 121 'The inside air is safe' 'Yes' 'No'

121 1, 0.00192, 1., 122 122,' Take inside air samples and test' 'Yes' 'No'

L 121 1, 122 1

122 1, 0.0306, 1., 124 124,' Get the correct test result' 'Yes' 'No'

L 122 1, 124 1

124 1, 0, 1., 107 107,' Re-venting' ' Yes ' ' No '

A 124 1, 135 * 210

L 124 0, 133 1

107 1, 0.025, 1., 131 108,' Safety manager authorizes the inside work' 'Yes' 'No'

108 1, 0.05, 1., 131 131,’ Task proceeding’ ’Stop’ ’Proceeding without the permit’

L 108 1, 132 1

L 108 0, 133 1

131 1,0,1., 132 132, ’ T1 Occupation accident’ ’No’ ’Yes’

L 131 1, 135 1

L 131 1, 1001 1

A 131 1, 135 * 1001

132 1,0,1., 133 133, ’ T1 Process accident’ ’No’ ’Yes’

L 132 1, 135 1

L 132 1, 1002 1

A 132 1, 135 * 1001

133 1,0,1.,134 134, ’ T1 Delay’ ’No’ ’Yes’

L 133 1, 135 1

L 133 1, 1003 1

134 1,0,1.,135 135, ’ T1 Wrong Result’ ’No’ ’Yes’

L 134 1, 135 1

L 134 1, 1004 1

135 1,0,1., 201 201, ’ T1 Failure’ ’No’ ’Yes’

L 135 1, 1005 1

! Task 2 Construction team building scaffold 210 1, 0.0724,1.,201 211, ’ Testing before enter the tank using multi-gas detector ’ ’Yes’ ’No’

L 210 0, 233 1

211 1, 0.00768,1.,212 231, ’ Enter the tank wearing oxygen aid and toxic-prevent equipment’ ’Yes’ ’No’ ! Occupation accident

L 211 1, 231 1

212 1, 0.00768,1.,201 232, ’ Enter the tank wearing electrostatic-prevent cloth and shoes’ ’Yes’ ’No’ ! Process accident

L 212 1, 232 1

201 1, 0.05, 1., 202 213, ’ building inside for less than 15 min and in turn’ ’ Yes’ ’ No’

L 201 1, 213 1

213 1, 0, 1., 202 220, ’Low O2’ ’ No’ ’ Yes’

220 1, 0.0001, 1., 221 221, ’ O2 gas alarm warning’ ’ Yes’ ’ No’

L 220 1, 222 1

221 1, 0.000128, 1., 202 222, ’ Get out of the tank in time’ ’ Yes’ ’ No’

L 221 1, 222 1

222 1, 0, 1., 202 231, ’Suffocation ’ ’No’ ’Yes’

L 222 1, 231 1

202 1, 0.000128, 1., 203 223, ’Object fall down or Personnel Fall’ ’No’ ’Yes’

L 202 1, 231 1

L 202 1, 233 1

223 1, 0.00768, 1., 231 231, ’Wearing safety helmet and rope’ ’Yes’ ’No’

L 223 1, 231 1

203 1, 0.000768, 1., 231 224, ’Build scaffold stable’ ’Yes’ ’No’ ! delay

L 203 1, 233 1

224 1, 0.05, 1., 231 225, ’Self check the scaffold’ ’Yes’ ’No’

225 1, 0.05, 1., 231 231, ’Authority check the scaffold’ ’Yes’ ’No’

L 225 1, 301 1

231 1,0,1., 232 232, ’ T2 Occupation accident’ ’No’ ’Yes’

L 231 1, 235 1

L 231 1, 1001 1

A 231 1, 235 * 1001

232 1,0,1., 233 233, ’T2 Process accident’ ’No’ ’Yes’

L 232 1, 235 1

L 232 1, 1002 1

A 232 1, 235 * 1001

233 1,0,1.,234 234, ’ T2 Delay’ ’No’ ’Yes’

L 233 1, 235 1

L 233 1, 1003 1

234 1,0,1.,235 235, ’ T2 Wrong Result’ ’No’ ’Yes’

L 234 1, 235 1

L 234 1, 1004 1

235 1,0,1., 301 301, ’ T2 Failure’ ’No’ ’Yes’

L 235 1, 1005 1

! Task 3 Inspection team Polishing Weld Joints

301 1,0,1., 302 302, ’Climbing up the scaffold’ ’Yes’ ’Fall’

L 301 1, 331 1

302 1,0.05,1., 303 310, ’Polishing for less than 15min in turn with sander tool’ ’Yes’ ’No’

L 302 1, 310 1

310 1, 0, 1., 303 320, ’Low O2’ ’ No’ ’ Yes’

320 1, 0.0001, 1., 321 321, ’ O2 gas alarm warning’ ’ Yes’ ’ No’

L 320 1, 322 1

321 1, 0.000064, 1., 303 322, ’ Get out of the tank in time’ ’ Yes’ ’ No’

L 321 1, 322 1

322 1, 0, 1., 303 331, ’Suffocation ’ ’No’ ’Yes’

L 322 1, 331 1

303 1, 0.000064, 1., 304 323, ’Object fall down or Personnel Fall’ ’No’ ’Yes’

L 303 1, 331 1

L 303 1, 333 1

323 1, 0.00384, 1., 331 331, 'Wearing safety helmet and rope' 'Yes' 'No'

L 323 1, 331 1

304 1,0.000384 , 1., 305 331, 'Spark from polish hurt eyes' 'No' 'Yes'

L 304 1, 331 1

305 1, 0.000384, 1., 331 306, 'Polished weld joint qualified' ' Yes ' 'No'

L 305 1, 333 1

306 1,0.05 , 1., 307 307, 'Check Polished weld joint qualified' ' Yes ' 'No'

L 306 1, 334 1

307 1, 0.000384,1., 331 331, 'Get hurt by sander tool' 'No' 'Yes'

L 307 1, 331 1

331 1,0,1., 332 332, ' T3 Occupation accident' 'No' 'Yes'

L 331 1, 335 1

L 331 1, 1001 1

A 331 1, 335 * 1001

332 1,0,1., 333 333, 'T3 Process accident' 'No' 'Yes'

L 332 1, 335 1

L 332 1, 1002 1

A 332 1, 335 * 1001

333 1,0,1.,334 334, ' T3 Delay' 'No' 'Yes'

L 333 1, 335 1

L 333 1, 1003 1

334 1,0,1.,335 335, ' T3 Wrong Result' 'No' 'Yes'

L 334 1, 335 1

L 334 1, 1004 1

335 1,0,1., 401 401, ' T3 Failure' 'No' 'Yes'

L 335 1, 1005 1

! Task 4 Inspection team performs the magnetic inspection

401 1,0,1., 406 406, ’Climbing up the scaffold’ ’Yes’ ’Fall’

406 1,0.05,1., 402 410, ’Inspection for less than 15min and in turn’ ’Yes’ ’No’

L 406 1, 410 1

410 1, 0, 1., 403 420, ’Low O2’ ’ No’ ’ Yes’

420 1, 0.001, 1., 421 422, ’ O2 gas alarm warning’ ’ Yes’ ’ No’

L 420 1, 422 1

421 1, 0.000064, 1., 402 422, ’ Get out of the tank in time’ ’ Yes’ ’ No’

L 421 1, 422 1

422 1, 0, 1., 403 431, ’Suffocation ’ ’No’ ’Yes’

L 422 1, 431 1

403 1,0.000384,1., 402 402, ’ Magnetic with the instrument according to the standard posture and time’ ’No’ ’Yes’

L 403 1, 434 1

402 1,0.000064,1., 404 404, ’ Spray fluorescence magnetic liquid towards the weld joints’ ’Yes’ ’No’

L 402 1, 434 1

404 1,0.000384,1., 405 405, ’ Irradiate the weld with ultraviolet rays in time’ ’No’ ’Yes’

L 404 1, 434 1

405 1,0.050064,1., 407 407, ’ Identify and Mark the crack correctly’ ’Yes’ ’No’

L 405 1, 434 1

407 1, 0.000064, 1., 408 423, ’Object fall down or Personnel Fall’ ’No’ ’Yes’

L 407 1, 431 1

L 407 1, 433 1

423 1, 0.00384, 1., 431 431, ’Wearing safety helmet and rope’ ’Yes’ ’No’

L 423 1, 431 1

408 1, 0.000384 , 1., 431 431, ' ultraviolet rays hurt eyes' 'No' 'Yes'

L 408 1, 431 1

431 1,0,1., 432 432, ' T4 Occupation accident' 'No' 'Yes'

L 431 1, 435 1

L 431 1, 1001 1

A 431 1, 435 * 1001

432 1,0,1., 433 433, 'T4 Process accident' 'No' 'Yes'

L 432 1, 435 1

L 432 1, 1002 1

A 432 1, 435 * 1001

433 1,0,1.,434 434, ' T4 Delay' 'No' 'Yes'

L 433 1, 435 1

L 433 1, 1003 1

434 1,0,1.,435 435, ' T4 Wrong Result' 'No' 'Yes'

L 434 1, 435 1

L 434 1, 1004 1

435 1,0,1., 1001 1001, ' T4 Failure' 'No' 'Yes'

L 435 1, 1005 1

1001 1,0,1., 1002 1002, 'Occupation accident' 'No' 'Yes'

1002 1,0,1., 1003 1003, 'Process accident' 'No' 'Yes'

1003 1,0,1., 1004 1004, 'Delay' 'No' 'Yes'

1004 1,0,1., 1005 1005, 'Wrong Result' 'No' 'Yes'

1005 1,0,1., 0 0, 'Failure ' 'No' 'Yes'

B.4 Input file of IDDA for HRT process

! Task 1 Prepare and Authorise the Work

101 1, 0, 1., 102 132,' The spherical tank inside is empty and purged ' ' Yes' 'No'

L 101 1, 132 1

102 1, 0.00192, 1., 103 110,' Insert the blind disc ' 'Yes' 'No'

110 1,0.025,1., 103 111, 'Recovery' ' Yes ' 'No '

L 110 1, 111, 1

111 1, 0, 1., 103 103,'Leak LPG into the spherical ' 'No' ' Yes '

A 111 1, 103 115 115

A 111 1, 105 120 120

L 111 1, 120 1

103 1, 0, 1., 104 104,'Open the manhole ' 'Yes' 'No'

115 1,0.00768,1., 116 116, 'Ignition source' 'No' 'Yes'

L 115 1, 116 1

116 1,0,1., 104 131, 'Flash Fire burnt the operator' 'No' 'Yes'

L 116 1, 131 1

104 1, 0.000032, 1., 105 131,'Set up electricity environment get shocked ' ' No ' ' Yes '

L 104 1, 131 1

105 1, 0.000192, 1., 106 106,' Venting with venting machine for more than 3 days' 'Yes' 'No'

L 105 1, 106 1

120 1,0,1., 106 132 'LPG Release' 'No' 'Yes'

L 120 1, 132 1

106 1, 0, 1., 107 121 'The inside air is safe' 'Yes' 'No'

121 1, 0.00192, 1., 122 122,' Take inside air samples and test' 'Yes' 'No'

L 121 1, 122 1

122 1, 0.0306, 1., 124 124,’ Get the correct test result’ ’Yes’ ’No’

L 122 1, 124 1

124 1, 0, 1., 107 107,’ Re-venting’ ’ Yes ’ ’ No ’

A 124 1, 135 * 210

L 124 0, 133 1

107 1, 0.025, 1., 131 108,’ Safety manager authorizes the inside work’ ’Yes’ ’No’

108 1, 0.05, 1., 131 131,’ Task proceeding’ ’Stop’ ’Proceeding without the permit’

L 108 1, 132 1

L 108 0, 133 1

131 1,0,1., 132 132, ’ T1 Occupation accident’ ’No’ ’Yes’

L 131 1, 135 1

L 131 1, 1001 1

A 131 1, 135 * 1000

132 1,0,1., 133 133, ’ T1 Process accident’ ’No’ ’Yes’

L 132 1, 135 1

L 132 1, 1002 1

A 132 1, 135 * 1000

133 1,0,1.,134 134, ’ T1 Delay’ ’No’ ’Yes’

L 133 1, 135 1

L 133 1, 1003 1

134 1,0,1.,135 135, ’ T1 Wrong Result’ ’No’ ’Yes’

L 134 1, 135 1

L 134 1, 1004 1

135 1,0,1., 201 201, ’ T1 Failure’ ’No’ ’Yes’

L 135 1, 1005 1

! Task 2 Inspection team Polishing Weld Joints with robot

210 1, 0.17,1.,201 211, ' Testing before enter the tank using multi-gas detector ' 'Yes' 'No'

211 1, 0.024,1.,212 231, ' Enter the tank wearing oxygen aid and toxic-prevent equipment' 'Yes' 'No' ! Occupation accident L 211 1, 231 1

212 1, 0.024,1.,201 232, ' Enter the tank wearing electrostatic-prevent cloth and shoes' 'Yes' 'No' ! Process accident L 212 1, 232 1

201 1,0.17,1., 202 202, 'Testing the robot with sample weld joint' 'Yes' 'No' L 201 1, 234 1

202 1,0.0004,1., 206 206, 'Locate the p-robot on the weld joint' 'Yes' 'No' L 202 1, 233 1

206 1,0.00000132,1., 207 216, 'p-robot move along the joints' 'Yes' 'No' L 206 1, 233 1

216 1, 0.07740032,1., 207 207 , 'recovery the situation outside the tank' 'Yes' 'No' L 216 1, 234 1

207 1,0.00000102 , 1., 208 219, 'Polished weld joint qualified' ' Yes ' 'No'

219 1,0.1, 1., 208 208, 'Check Polished weld joint qualified' ' Yes ' 'No'

L 219 0, 233 1

L 219 1, 234 1

208 1,0.0024 , 1., 220 220, 'Clean the iron pieces in time' ' Yes ' 'No'

L 208 1, 220 1

220 1,0.00000002 , 1., 230 221, 'p-robot fall down' ' Yes ' 'No'

221 1,0.0024, 1., 230 230, 'the p-robot locked with protect rope' 'Yes' 'No'

L 221 1, 230 1

230 1,0,1., 231 231, ' T2 robot or tank damaged' 'No' 'Yes'

L 230 1, 235 1

L 230 1, 1000 1

A 230 1, 235 * 1000

231 1,0,1., 232 232, ' T2 Occupation accident' 'No' 'Yes'

L 231 1, 235 1

L 231 1, 1001 1

A 231 1, 235 * 1000

232 1,0,1., 233 233, 'T2 Process accident' 'No' 'Yes'

L 232 1, 235 1

L 232 1, 1002 1

A 232 1, 235 * 1000

233 1,0,1.,234 234, ' T2 Delay' 'No' 'Yes'

L 233 1, 235 1

L 233 1, 1003 1

234 1,0,1.,235 235, ' T2 Wrong Result' 'No' 'Yes'

L 234 1, 235 1

L 234 1, 1004 1

235 1,0,1., 301 301, ' T2 Failure' 'No' 'Yes'

L 235 1, 1005 1

! Task 3 Inspection team performs magnetic inspection with m-robot

301 1,0.17,1., 302 302, 'Testing the sensitivity of m-robot with a sample' 'Yes' 'No'

L 301 1, 334 1

302 1,0.0004,1., 304 304, 'Locate the m-robot on the weld joint' 'Yes' 'No'

L 302 1, 333 1

304 1,0.00000132,1., 305 321, 'M-robot move along the joints' 'Yes' 'No'

L 304 1, 334 1

321 1, 0.07740032,1., 305 305 , 'recovery the situation outside the tank' 'Yes' 'No'

L 321 1, 334 1

305 1,0.00000154,1., 306 306, 'Spray the fluorescence magnetic liquid' 'Yes' 'No'

L 305 1, 334 1

306 1,0.01,1., 307 308, 'Detect and Judge as a crack' 'Correctly' 'Falsely'

307 1,0.00000032,1., 308 310, 'Stop and send an alarm' 'Yes' 'No'

L 307 1, 334 1

308 1,0.1,1., 310 310, 'Confirm the crack' 'Yes' 'No'

L 308 1, 334 1

310 1,0.00000002 , 1., 326 326, 'M-robot fall down' ' Yes ' 'No'

326 1,0.0024, 1., 330 330, 'the M-robot locked with protect rope' 'Yes' 'No'

L 326 1, 330 1

330 1,0,1., 331 331, ' robot damage' 'No' 'Yes'

L 331 1, 335 1

L 331 1, 1000 1

331 1,0,1., 332 332, ' T3 Occupation accident' 'No' 'Yes'

L 331 1, 335 1

L 331 1, 1001 1

332 1,0,1., 333 333, 'T3 Process accident' 'No' 'Yes'

L 332 1, 335 1

L 332 1, 1002 1

333 1,0,1.,334 334, ' T3 Delay' 'No' 'Yes'

L 333 1, 335 1

L 333 1, 1003 1

334 1,0,1.,335 335, ' T3 Wrong Result' 'No' 'Yes'

L 334 1, 335 1

L 334 1, 1004 1

335 1,0,1., 1000 1000, ' T3 Failure' 'No' 'Yes'

L 335 1, 1005 1

1000 1,0,1., 1001 1001, ’Robot damaged’ ’No’ ’Yes’

1001 1,0,1., 1002 1002, ’Occupation accident’ ’No’ ’Yes’

1002 1,0,1., 1003 1003, ’Process accident’ ’No’ ’Yes’

1003 1,0,1., 1004 1004, ’Delay’ ’No’ ’Yes’

1004 1,0,1., 1005 1005, ’Wrong Result’ ’No’ ’Yes’

1005 1,0,1., 0 0, ’Failure ’ ’No’ ’Yes’