Journal Pre-proof

Systemic Risk Analysis of Complex Socio-Technical Systems from the Safety-II Perspective

Massoud Mohsendokht , Huanhuan Li , Christos Kontovas , Chia-Hsun Chang , Zhuohua Qu , Zaili Yang

PII: S0951-8320(26)00016-5
DOI: https://doi.org/10.1016/j.ress.2026.112200
Reference: RESS 112200

To appear in: *Reliability Engineering and System Safety*

Received date: 15 September 2025
Revised date: 12 December 2025
Accepted date: 5 January 2026

Please cite this article as: Massoud Mohsendokht , Huanhuan Li , Christos Kontovas , Chia-Hsun Chang , Zhuohua Qu , Zaili Yang , Systemic Risk Analysis of Complex Socio-Technical Systems from the Safety-II Perspective, *Reliability Engineering and System Safety* (2026), doi: https://doi.org/10.1016/j.ress.2026.112200

Systemic Risk Analysis of Complex Socio-Technical Systems from the Safety-II Perspective

Massoud Mohsendokht [a], Huanhuan Li [a], Christos Kontovas [a], Chia-Hsun Chang [a], Zhuohua Qu [b], Zaili Yang [a]*

[a] *Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK*

[b] *Liverpool Business School, Liverpool John Moores University, United Kingdom*

*Corresponding author: z.yang@ljmu.ac.uk

ABSTRACT

Modern complex socio-technical systems demand systemic risk analysis approaches that can holistically address the interdependencies between human, technological, and organizational components. Traditional models often fall short in capturing the dynamic and emergent nature of these interactions. This study introduces a novel, integrated risk analysis framework grounded in the Safety-II paradigm, which emphasizes understanding how systems succeed under varying conditions rather than focusing solely on failure. The proposed methodology combines the Functional Resonance Analysis Method (FRAM) with Bayesian Networks to overcome FRAM's qualitative limitations and enable quantitative assessment of performance variability. The framework is further enriched by integrating complementary techniques, including Monte Carlo Simulation and canonical probabilistic models. This holistic toolkit enables a rigorous and scalable approach for modelling uncertainty and systemic variability across complex operational environments. The methodology is demonstrated through a case study of seaport operations, a representative example of a complex socio-technical system. The results show that the integrated Safety-II-informed framework improves the quantification of systemic risk and enhances the capacity to manage complexity and uncertainty in real-world settings.

Keywords: Systemic risk, Complex socio-technical system, Safety II, FRAM, Bayesian Network, Seaport Operations

1. INTRODUCTION

Complex Socio-Technical Systems (CSTS) are defined by tightly interconnected structures, unpredictable workflows, non-linear operations, and intricate interactions among their elements. These systems encompass the interplay of human, technological, and environmental factors within an organizational context (Baxter and Sommerville, 2011; Bayramova et al., 2023; Jensen and Aven, 2018). Traditional risk analysis methods, such as fault tree analysis, event tree analysis, and probabilistic safety assessment, are primarily grounded in the Safety-I paradigm. These approaches operate on several foundational assumptions: systems can be decomposed into simpler components; their functioning is categorized as either successful or failed; risk analysis depends on predefined cause-and-effect relationships; and event sequences are assumed to be linear. While this methodology proved effective for purely technological systems and was widely applied in critical industries such as chemical, nuclear, and aviation during the 20th century, its limitations became apparent when dealing with CSTS *(Aven, 2022; Mohsendokht and Jamshidi, 2021)*.

Safety-I philosophy, rooted in traditional thinking, struggles to accommodate the dynamic, nonlinear, and emergent nature of modern systems, making its continued application in the 21st century increasingly questionable *(Hollnagel et al., 2015)*. To address these limitations, a new paradigm known as Safety-II has emerged. Rather than focusing solely on the prevention of failures, Safety-II emphasizes ensuring that "as many things as possible go right." This approach adopts a proactive stance, recognizing the adaptability of human operators and underscoring the importance of monitoring everyday performance variability as a means of sustaining system safety.

Over the past decade, this paradigm has sparked extensive discussion among researchers, who have both supported and critiqued its underlying philosophy; a detailed exploration of which lies beyond the scope of this study *(Aven, 2022; Cooper, 2022; Hollnagel, 2018; Martinetti et al., 2019; Provan et al., 2020)*. However, the research trend indicates that the Safety-II concept has gained significant traction, with scholars from various disciplines incorporating it into their studies. Applications span diverse fields, including maritime operations *(Adhita et al., 2023; Qiao et al., 2021; Wahl et al., 2020)*, energy systems *(Riemersma et al., 2024)*, aviation *(Yang et al., 2017)*, chemical industry *(Yu et al., 2020)*, construction *(Martinetti et al., 2019; Martins et al., 2022)*, transportation *(Papadimitriou et al., 2022; Wang et al., 2020)*, and nuclear power plants *(Ham and Park, 2020; Park et al., 2018)*.

Despite the growing conceptual appeal of the Safety-II concept across domains, efforts to translate its principles into repeatable, decision-oriented analyses for CSTS remain fragmented. Existing operationalisations tend to be either qualitative (e.g., mapping work-as-done, identifying functional dependencies, and general recommendations to improve safety) or narrowly quantitative (e.g., indicator scoring or isolated simulations), often without a formal mechanism to represent everyday performance variability, propagate its effects through interdependent functions, and address uncertainty in a transparent way.

Recent advances over the past years have sought to address this gap by proposing semi-quantitative and quantitative approaches in which Functional Resonance Analysis Method (FRAM) serves as the central modelling framework. These efforts include the use of Monte Carlo sampling and explicit propagation rules to characterise upstream-downstream variability within FRAM models *(Kim and Yoon, 2021; Patriarca et al., 2017)*, as well as the integration of FRAM with BNs or dynamic BNs *(Peng et al., 2023; Zarei et al., 2022)*. Such combinations provide a principled calculus for fusing heterogeneous evidence, updating beliefs, and supporting diagnostic and prognostic reasoning in interdependent systems. Collectively, these developments have paved the way toward more rigorous operationalisation of the Safety-II concept in CSTS. However, existing studies still have limitations, in which some concentrate on modelling functional interactions while overlooking internal and external variabilities within individual functions, whereas others emphasise these variabilities but fail to capture the dynamic interplay between functions.

To address these gaps, this paper proposes an integrated framework for systemic risk analysis within the context of CSTS, aligning with the principles of the Safety-II concept. The novel methodology integrates FRAM and BN with advanced analytical tools, including Monte Carlo Simulation (MCS), canonical probabilistic methods, Dempster-Shafer theory, and criticality matrix. The key contributions of this study can be summarized as follows:

1) Comprehensive analysis of CSTS elements: Each element of the CSTS, including technological, human, and organizational functions, is analysed to assess their internal and external performance variabilities. These variabilities account for factors such as operational uncertainties, environmental conditions, and human performance fluctuations.
2) Interaction between functions: The interactions among related functions are systematically examined to identify and track their upstream-downstream performance variability. This includes assessing their potential impacts, either negative, damping, or even positive on the entire system. Such an analysis helps to highlight critical dependencies and emergent behaviours within the system.
3) Retrospective and prospective risk analysis: The proposed framework enables both retrospective and prospective evaluations of the performance variability. This dual perspective equips decision-makers with actionable insights to address risks effectively.
4) Support for risk-based decision-making: By quantifying and visualizing variabilities across the CSTS, the framework empowers decision-makers to prioritize interventions and implement targeted measures to manage identified risks.

The remainder of this paper is structured as follows: Section 2 provides a concise literature review on CSTS, the application of the Safety-II concept, outlining the methodologies currently applied in CSTS and highlighting the identified research gaps. Section 3 presents a detailed discussion of the adopted methodology, with an in-depth explanation of the various techniques employed. Section 4 demonstrates the application of the proposed methodology to seaport operations and includes a discussion of the results and their interpretation. Finally, Section 5 summarizes the key insights derived from this study and formulates the conclusions.

2. LITERATURE REVIEW

2.1. Systemic risk analysis of complex socio-technical systems

As previously noted, CSTS are networked configurations of individuals, technologies, rules, and environments whose behaviour emerges from numerous non-linear interactions rather than from any single component. Within such systems, accidents frequently originate from mismatches, tight couplings, and feedback across organisational, human, and technical layers, rather than solely from component failure or human error. Systemic risk analysis perspectives emphasise that safety performance depends on how constraints are specified, implemented, and monitored throughout the entire system structure, and that adverse outcomes may arise from otherwise normal local variability when influenced by goal conflicts and resource pressures *(Leveson, 2004)*.

This inherent complexity underscores the relevance of the Safety-II perspective, which focuses on understanding how work typically succeeds despite performance variability. Safety-II recognises that the same adaptations that enable successful outcomes can, under certain circumstances, interact to produce failure. By shifting the analytical focus toward everyday performance, operational trade-offs, and resilience capacities, Safety-II provides a more robust foundation for systemic risk analysis and for designing systems that remain tolerant and adaptive in the face of variability *(Provan et al., 2020)*.

In this regard, efforts have been made to introduce techniques for systemic risk analysis, including the Function Analysis System Technique (FAST), the Structured Analysis and Design Technique (SADT), the Systematic Human Error Reduction and Prediction Approach (SHERPA), the Accident Causation Analysis and Taxonomy (ACAT), the Systems Theoretic Accident Model and Processes (STAMP), and the FRAM. Table 1 presents a concise comparative analysis of these methods, highlighting their respective strengths and limitations in the context of CSTS risk analysis.

Table 1. Comparative Overview of Systemic Risk Analysis Methods.

Method	Analytical Focus / Application Domain	Advantages	Limitations	Key References
FAST	Employed primarily in engineering design and value analysis to map functional logic ("how" and "why") between system elements.	- Promotes function-oriented rather than component-based thinking. - Facilitates stakeholder engagement and creative exploration of alternatives.	- Outcomes are highly dependent on facilitator expertise. - Provides only static representations; limited capacity to model temporal or emergent behaviours.	*(Bytheway, 2007)*
SADT	Utilized for hierarchical decomposition of system functions, specifying inputs, outputs, controls, and mechanisms in structured diagrams.	- Offers a standardized, formalized framework for system modelling and documentation. - Enhances communication among interdisciplinary teams.	- Inflexible in dynamic or rapidly evolving environments. - Lacks constructs for sequencing or dynamic interactions; models may become complex and unwieldy.	*(Ahmed et al., 2014)*
SHERPA	Designed to identify and classify potential human errors based on task analysis within complex systems.	- Systematic prediction of error modes with direct links to remedial actions. - Supports targeted safety interventions and human factors engineering	- Requires comprehensive task decomposition in advance. - Resource-intensive and time-consuming to implement at scale.	*(Stanton, 2004)* *(Catelani et al., 2021)*
ACAT	Focused on classifying and analysing causal factors in past accidents across technical, human, managerial, and environmental	- Provides a comprehensive taxonomy for multi-dimensional accident analysis.	- Primarily retrospective in nature; limited use for prospective risk assessment. - Requires domain-specific adaptation.	*(Li et al., 2017)*

	domains.	- Enables systematic tracing of failure paths		
STAMP	Conceptualizes accidents as results of inadequate control and feedback within socio-technical systems, grounded in systems theory.	- Captures unsafe interactions and emergent risks in complex adaptive systems. - Supports high-level modelling of safety constraints and control structures	- Requires substantial conceptual understanding and modelling effort. - Predominantly qualitative; quantitative extensions remain limited	*(Ceylan et al., 2021)* *(Sun et al., 2022)*
FRAM	Models complex socio-technical systems by examining how functional performance variability propagates and interacts over time.	- Well-aligned with the Safety-II paradigm; captures both successful and adverse outcomes. - Explicitly models interactions among human, technical, and organizational elements. - Enables dynamic assessment of functional resonance and emergent risks.	- Model development is resource-intensive and requires deep system understanding. - Quantitative applications are evolving but not yet standardized.	*(Erik, 2017)* *(Yu et al., 2024)*

2.2. FRAM application

Among the above-mentioned techniques, the FRAM has gained significant popularity for systemic risk analysis in CSTS due to several compelling advantages. Unlike traditional methods, FRAM avoids decomposing systems into individual components and operates independently of cause-effect analysis, aligning seamlessly with the principles of the Safety-II paradigm. Furthermore, it identifies the various elements of a CSTS (Human, technological, and organizational factors) and addresses them holistically while accounting for their interactions and interdependencies. Additionally, FRAM enables detailed monitoring and analysis of the performance variability of each function, its influence on downstream functions, and its overall impact on the entire system. FRAM models CSTS by focusing on the functions that describe what the system does, rather than its physical components or organisational structure. Each function is depicted as a hexagon with six aspects, including Input, Output, Preconditions, Resources, Control, and Time that define its behaviour and interaction with other functions. Couplings between functions are represented by arrows, indicating how the output of one function can influence the input, control, or resource requirements of another. Variability emerging in any function may propagate through these couplings, and this functional interaction structure forms the basis for analysing the performance variability.

The four main principles of FRAM can be summarized as follows (Erik, 2017): First, the nature of success and failure is equivalent; in other words, everyday work variability determines whether outcomes are positive or negative. Second, individuals and organizations must make ongoing, often approximate, adjustments to adapt to changing conditions. Third, outcomes both positive and negative of a CSTS emerge from interactions among multiple system functions rather than from individual components alone, meaning outcomes cannot be traced directly to isolated causes. Fourth, functional resonance refers to the amplification of normal function variability due to unexpected interactions. It emphasizes the importance of identifying areas where such resonances may occur, as they can lead to significant system-wide consequences.

Despite its many advantages, the FRAM remains primarily a qualitative approach, lacking the capability to provide quantitative measurements for interpreting performance variability. This limitation is widely recognized as a significant drawback. To address this issue, researchers have investigated various approaches to enhancing FRAM by incorporating standardized and quantitative techniques. One of the earliest efforts in this direction was undertaken by Rosa et al. *(Rosa et al., 2015)*, who combined FRAM with the Analytical Hierarchy Process (AHP) to generate numerical rankings. Patriarca et al. *(Patriarca et al., 2017)* introduced an innovative semi-quantitative FRAM-based approach by integrating it with MCS, enabling the representation of performance variability as discrete probability distributions. The integration of fuzzy logic theory with FRAM has also been proposed in multiple studies, offering another pathway to quantification *(Hirose and Sawaragi, 2020, 2019; Slim and Nadeau, 2020)*. In their work, Lee and Chung *(Lee and Chung, 2018)* developed a method to quantify Human-System Interaction (HSI) variability and assess criticality using a semi-quantitative FRAM process.

More advanced techniques have emerged in recent years, including the integration of machine learning and data-driven approaches with FRAM, which have been applied across various domains. BNs have also been explored as a powerful probabilistic tool for quantifying FRAM. For instance, Zarei et al. *(Zarei et al., 2022)* developed a causation model based on FRAM, which they incorporated into a dynamic BN to analyse internal and external performance variability, referred to as uncoupled variability, within the petrochemical industry. In maritime operations, Guo et al. *(Guo et al., 2023)* proposed a similar approach, further enhanced by embedding a Markov model to analyse the evolution of collision risk during ship pilotage. These advancements demonstrate the growing efforts to integrate qualitative and quantitative analyses in FRAM applications *(Wang et al., 2025; X. Yang et al., 2025; Zhao et al., 2025)*.

2.3. Research gaps

Following a comprehensive review of the current literature, a recurring critique highlights the lack of a systematic framework for improving safety performance that effectively integrates both qualitative and quantitative approaches. Qualitative approaches, while valuable for conceptual exploration, frequently lack systematic and quantifiable measures. These approaches often focus on describing the contrast between work-as-imagined and work-as-done, mapping functional dependencies, and providing general recommendations for improvement, but rarely progress toward measurable, evidence-based interventions.

Quantitative approaches, on the other hand, also exhibit notable limitations. Some studies prioritise modelling functional interactions but neglect internal and external variabilities within individual functions. Others emphasise characterising such variabilities yet fail to adequately capture the dynamic interdependencies among functions. In the first case, functions are frequently treated as nodes with fixed or weakly varying parameters, under-representing internal variability (e.g., workload, expertise drift, equipment degradation) and external variability (e.g., demand surges, environmental conditions, regulatory or organisational changes). As a result, resonance pathways are computed over unrealistically stable functions, with uncertainty addressed through ad hoc sensitivity ranges rather than systematic propagation techniques. In the second case, although intra-function variability is richly characterised, functional couplings are simplified or omitted. Outputs are often aggregated into single indices, temporal dynamics are suppressed, and dependencies are assumed independent. This prevents the transmission of cross-scale feedback, buffering effects, and transient accumulations through the functional network. Additionally, the use of static or scenario-specific parameterisation and limited evidence fusion or validation further restricts robust prospective "what-if" analysis.

A more holistic methodology is therefore required which could retain FRAM's functional topology, embeds stochastic and state-dependent models for each function, and employs a probabilistic propagation engine to fully operationalise the principles of Safety-II in CSTS. The methodology proposed in this study seeks to address these gaps, as detailed in the following sections.

3. METHODOLOGY

This section proposes a novel systemic risk analysis methodology based on a hybrid approach combining FRAM and BN, representing three key elements of CSTS: technological, human, and organizational functions. FRAM is utilized to describe the complex interrelationships among various functions, while BN enables the quantitative analysis of this complexity. Figure 1 illustrates the overall methodology, structured into four consecutive phases.

- Phase 1: Based on Hierarchical Task Analysis (HTA) and the principles of FRAM, the functions, associated variabilities, and couplings between functions are identified, leading to the construction of the final FRAM model.
- Phase 2: Each function is represented as either a technological, human, or organizational function. The internal variability within each function is modelled using a BN, in which the interrelationships among its internal contributing factors are defined both qualitatively and quantitatively.
- Phase 3: The FRAM, serving as the primary model, is integrated with the BN to represent variability, incorporating prior probabilities, conditional probability tables, and model validation.
- Phase 4: The model is interpreted by identifying resonances, whether negative or damping, recognizing critical functions through monitoring interactions between them, detecting resonant patterns, and ultimately extracting insights and implications.

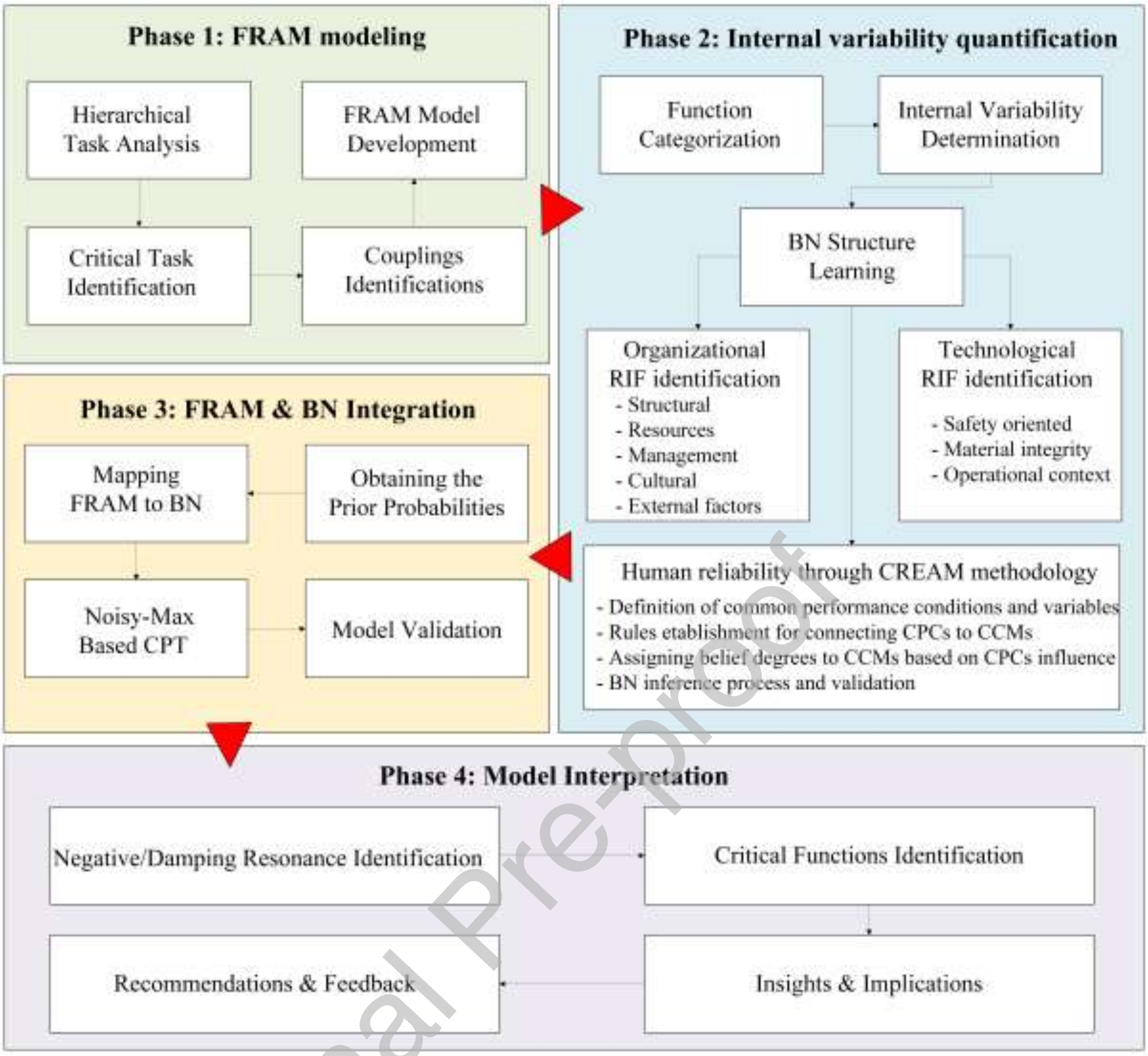

Figure 1: The Conceptual structure of the developed methodology.

3.1. FRAM modelling

In the first phase, an HTA is developed to better understand the activities within the process under study and to provide a general overview of its tasks. The hierarchical structure of HTA enables detailed analysis of specific tasks and helps clarify the relationships among them. HTA has been extensively described in prior research (Salmon et al., 2010; Stanton, 2006). Once the HTA is developed, key functions are identified and selected for further analysis through FRAM modelling. FRAM is employed to qualitatively analyse the effective operation of a CSTS.

Based on the principles of FRAM discussed earlier, the model can be constructed by the following steps outlined below:

1) Identification of functions: The results from the HTA inform the FRAM construction. Activities that significantly contribute to the overall process are identified as candidate functions.
2) Definition of aspects: Each function is characterized by six aspects: input, output, resource, pre-condition, control, and time.

3) Determination of couplings: By understanding the flow of information or resources within the system, links between different aspects of various functions are identified, allowing for visualization of interdependencies among functions.
4) Identification of variabilities: Function variability refers to deviations in a function's output caused by factors from internal, external, or upstream functions.

Once the FRAM structure is constructed, each function can be characterized by potential performance variabilities. In FRAM modelling, three types of variability are considered: (1) Internal variability: which originates from factors within the function itself, such as staff training levels and equipment maintenance schedules; (2) External variability: which is driven by external factors like weather conditions, geopolitical events, market demands, and security issues; (3) Upstream Variability Index (UVI), which captures the effects of interdependencies with upstream functions that affect downstream functions, such as the impact of container unloading efficiency and speed on the subsequent transport accuracy and timing to yard storage in a seaport. In this paper, the first two variabilities are referred to as Self-contained Variability Index (SVI), which pertains to performance fluctuations caused by internal and external factors that do not arise from interactions between system functions.

These variabilities, interpreted as abnormal daily fluctuations, manifest in different ways, known as phenotypes, according to Erik (2017). Phenotypes may include aspects such as timing, precision, speed, distance, force, duration, and direction. Depending on the nature of the analysis, a suitable combination of these phenotypes is chosen for FRAM analysis. In this paper, timing and precision are selected to represent the performance variability of the functions. Timing represents the punctuality of activities being conducted. The output of a function may occur too early, on time, late, or in the worst case, missed which means it arrives too late to be useful for its intended purpose or is not produced at all (Kaya et al., 2021). Regarding precision, an output can be accurate, satisfactory, inaccurate, or, in the worst case, faulty. From a systemic perspective, performance variability arises from local adjustments made to meet performance demands and ensure the functioning of a CSTS. To conduct a meaningful analysis, it is essential to evaluate the potential variability of each function. Therefore, a unified representation of performance variability is needed, enabling an aggregated view across different types of functions. To this end, integrating these two phenotypes not only provides a unified language for describing performance variability among functions but also facilitates the interpretation of interactions between these functions *(Grabbe et al., 2022; Slim and Nadeau, 2020; Zinetullina et al., 2021)*. Table 2 presents the results of this integration using four qualitative scales: stable (ST), low variable (LV), moderately variable (MV), and highly variable (HV).

Table 2. Unification of performance variability based on time and precision phenotypes.

		Time			
		Early	Timely	Delayed	Missed
Precision	Accurate	ST	ST	LV	HV
	Satisfactory	LV	LV	MV	HV
	Inaccurate	MV	MV	HV	HV
	Faulty	HV	HV	HV	HV

In this context, "ST" performance is achieved when activities are both timely and accurate, indicating no variability and reliable outcomes. It is the only case where "work as imagined" corresponds exactly to "work as done". "LV" describes situations where performance may show slight deviations but remains satisfactory, being either timely or accurate. "MV" occurs when inaccuracies or delays begin to affect performance, though it remains functional. "HV" represents significant deviation, with outputs frequently delayed, missed, or faulty, leading to unreliability and potential system disruption.

3.2. BN modelling

For a quantitative analysis of FRAM, using BN to represent qualitative performance variability scales in a numerical form is highly effective. This approach offers two primary advantages. First, since performance variability has four defined states, BN can seamlessly manage these multi-state conditions, accommodating the complexity introduced by numerous interacting functions within a system. Second, performance variability can be expressed as probability percentages, a task well-suited to BN's strength in handling probabilistic analysis and uncertainty. Thus, integrating BN with FRAM enables a robust approach to systemic risk analysis in CSTS, leveraging probabilistic reasoning to capture the nuanced variability and interdependencies inherent in these environments. To begin, it is essential to differentiate functions based on their inherent characteristics, categorizing them into three primary types: technological functions, human functions, and organizational functions. Each category represents a distinct aspect of the system with unique dependencies, behaviours, and potential risks. Separate BN models are developed for each of these categories to capture the specific interactions, uncertainties, and causal relationships within each function type, a concept referred to as SVI.

In this respect, a structured pipeline was used to identify and justify priority nodes and states for each function: (i) literature-based scoping to enumerate candidate nodes and state options; (ii) expert review to apply inclusion/exclusion criteria and finalise observable, non-overlapping states; and (iii) validity checks via a BN-level sensitivity screening.

In the first step, candidates were compiled from a comprehensive review of published sources, retaining variables with plausible causal relevance to the target node. In the next step, a multidisciplinary domain expert panel merged or removed overlapping items, confirmed state labels, and standardised state counts to 2-3 for BN tractability. Inclusion/exclusion criteria are set as follows: operationalisability (observable in this context and discretisable into mutually exclusive, collectively exhaustive states), non-redundancy (no conceptual duplication), and interpretability (states understandable to practitioners). In the last step, the BN sensitivity screen fixed each parent to each of its states, recalculated the child’s probability, computed the parent swing (max-min change), and verified monotonicity (worsening states increased risk); nodes with negligible or non-monotone effects were revised or omitted.

3.2.1. Organizational functions

Organizational factors play a crucial role in system safety, either enhancing or impairing the safety performance of a CSTS. Within an organization, numerous interactions occur among various components, including staff, operators, management, structure, and culture, among others *(Li et al., 2012; Pence and Mohaghegh, 2020)*. To explicitly account for the impact of organizational factors on system performance variability and to capture the collective nature of its constituent elements, it is essential to consider all relevant aspects across multiple dimensions. These dimensions include social factors (e.g., safety culture, level of training), structural factors (e.g., authority gradients), resources (e.g., financial), management (e.g., leadership quality), and even external factors (e.g., geopolitical influences). Furthermore, the interactions among these dimensions must also be thoroughly examined *(Mohaghegh et al., 2009)*. Table 3 outlines the contributing factors of organizational functions, along with their sub-factors and corresponding descriptions, within the context of BN development. In this respect, efforts are made to define the states of each node to ensure an appropriate depth of causality in the model, while accounting for the objectives of systemic risk analysis and the multidimensional nature of organizational factors.

Table 3: The identified contributing factors to the performance variability of organizational functions.

Categories	Nodes	States	Descriptions	Reference
External factors	Regulation and enforcement	Strict, moderate, lax	Refers to the laws, regulations, standards, and oversight mechanisms established by governmental or regulatory bodies that an organization is required to follow. Stricter rules generally lead to improved organizational performance in the relevant functions.	*(Donelson et al., 2023)*
	Market conditions	Favourable, unfavourable	Refers to the various economic factors and dynamics that impact the supply and demand for goods and services within a specific market. Unfavourable market conditions can significantly influence an organization's operational decisions, strategic planning, and overall performance.	*(Germain et al., 2008)*
	External stakeholder relationships	Strong, average, weak	A strong relationship with external stakeholders can enhance organizational performance by fostering trust, facilitating resource access, and enabling smoother collaboration. Conversely, a weak relationship may lead to communication gaps, reduced support, and potential conflicts, leading to an increased performance variability.	(Hillman and Keim, 2001)
	Geopolitical factors	Stable, tense	Intense geopolitical factors, such as international conflicts, trade policies, tariffs, and economic sanctions, can negatively impact organizational performance, as managing these external pressures requires complex and challenging decision-making. In stable conditions, however, the organization is relieved from such difficulties.	*(Kuai and Wang, 2025)*
	Environmental factors	Favourable, unfavourable	Environmental factors, such as climate change and natural disasters, can disrupt operations, increase costs, and require	*(Busch, 2011)*

			investment in sustainable practices. Failure to respond, adapt, and recover effectively from these factors can damage the organization's reputation, hinder compliance, and negatively impact overall performance.	
	Security factors	Secure, insecure	Security factors, including data breaches, cyber threats, and physical security risks, can compromise sensitive information, disrupt business continuity, and increase the organizational performance variability.	*(Hasan et al., 2021)*
Organizational Structure	Span of control	Wide, balanced, narrow	Refers to the horizontal aspect of management, i.e., how many employees are directly under the supervision of a single manager. A wider span of control means fewer managers are needed, leading to a flatter organizational structure. A narrower span of control requires more managers, leading to a taller structure. A balanced span of control indicates of the appropriate number of managers.	*(Remenova et al., 2018)*
	Communication paths	Adequate, inadequate	Communication paths refer to the adequacy and quality of communication between different levels of an organization. When communication is sufficient and effective, the organization's performance variability becomes more stable.	*(Musheke and Phiri, 2021)*
	Authority gradient	Steep, balanced, shallow	An authority gradient describes the hierarchy of power within an organization, influencing how freely subordinates can challenge superiors. A steep gradient discourages lower-ranking individuals from speaking up, while a balanced gradient promotes open dialogue across levels. Conversely, a shallow gradient can lead to a chaotic environment.	*(Luva and Naweed, 2024)*
Organizational resources	Equipment resources	Adequate, inadequate	An adequate amount of equipment resources is essential for stable organizational performance.	*(Ozdemir et al., 2023)*
	Human resources	Adequate, inadequate	An adequate number of personnel is essential for stable organizational performance.	*(Natsir et al., 2024)*
	Financial resources	Adequate, inadequate	An adequate number of financial resources is essential for stable organizational performance.	*(Carmeli and Tishler, 2004)*
	Information resources	Adequate, inadequate	An adequate amount of information resources is essential for stable organizational performance.	*(Pashutan et al., 2022)*
	Time resources	Adequate, inadequate	An adequate amount of time resources is critical for meeting deadlines, maintaining productivity, and ensuring	*(Aeon et al., 2021)*

			efficient workflow.	
Organizational management	Resource management	Efficient, moderately efficient, inefficient	Refers to the organized efforts and procedures an organization implements to allocate existing resources effectively and efficiently.	*(Wang et al., 2022)*
	Leadership quality	Strong, moderate, weak	Refers to the effectiveness and characteristics of leaders within an organization. Strong and high-quality leadership is crucial for setting the direction, inspiring employees, and ensuring the achievement of organizational goals.	*(Nasim et al., 2023)*
	Communication effectiveness	Adequate, inadequate	Refers to the quality of communication within an organization and its impact on achieving stable performance. It encompasses the clarity, accuracy, and timeliness of information shared among team members. Clear communication promotes collaboration, minimizes misunderstandings, and aligns everyone with the organization's objectives, ultimately ensuring consistent performance.	*(Noor Arzahan et al., 2022)*
	Rules & regulations implementation	Compliant, partially compliant, incompliant	Refers to the effectiveness with which an organization enforces and adheres to internal policies, standards, and external regulations governing its operations. Greater compliance with these rules and regulations leads to more stable organizational performance variability.	*(Pedrosa et al., 2025)*
	Emergency management	Strong, moderate, weak	Refers to the organized efforts and procedures that an organization establishes strategies to handle emergencies by planning ahead, managing responses, and facilitating recovery efforts, including natural disasters, technological incidents, security threats, and other unexpected events that may disrupt normal operations. The stronger the emergency management, the more stable the organization's performance variability.	*(Mees et al., 2016)*
Organizational culture	Education/training	Adequate, inadequate	An adequate level of education and training among personnel contributes to a vibrant organizational culture.	*(RAHAMAN et al., 2023)*
	Information sharing	Adequate, inadequate	Refers to the process of exchanging relevant information including data, knowledge, insights, and updates among individuals, teams, departments, or organizations. Adequate level	*(Diem Le et al., 2023)*

			of information sharing is crucial for overall organizational efficiency.	
	Safety culture	Rich, moderate, poor	Refers to shared mindset, outlook, and priorities of employees concerning safety practices and standards within an organization. It encompasses how safety is prioritized, communicated, and practiced at all levels, from management to front-line workers. A rich safety culture fosters a proactive approach to managing risks.	*(Noor Arzahan et al., 2022)*
	Organizational cohesion	High, moderate, low	It reflects how well employees work together toward common goals, the strength of relationships within the organization, and the overall sense of belonging and loyalty that employees feel. An organization with high level of cohesion typically experiences higher levels of productivity, and performance stability.	*(Grossman et al., 2022)*
	Employee inclusivity	Inclusive, moderately inclusive, exclusive	Encompasses initiatives aimed at fostering an inclusive and supportive workplace where every employee feels respected, appreciated, and encouraged to actively participate. A high level of inclusivity within an organization fosters a rich organizational culture.	*(Chinenye Gbemisola Okatta et al., 2024)*

3.2.2. Technological functions

Technological functions are primarily driven by machinery, equipment, or software and represent automated processes or technical operations within CSTS. These functions rely on the technical features of the system to perform specific tasks. Technological functions are typically characterized by precision, consistency, and a predictable range of variability, usually governed by design specifications, technical capabilities, or programmed protocols. To determine the SVI of this function, the contributing factors to its performance variability must first be identified. Performance variability states, designated as the child node of the BN, include ST, LV, MV, and HV. The parent nodes, representing primary influences on performance variability, can be categorized into three main groups: safety-oriented factors, material integrity factors, and operational context factors. Safety-oriented factors encompass the protocols, practices, and resources dedicated to ensuring operational safety, reliability, and performance stability. These factors reflect the effectiveness of safety management within the system and play a crucial role in building resilience to variability and failure. Key contributors include maintenance activities, inspection policies, and reliability indices. Material integrity factors represent the physical condition and degradation of materials over time, accounting for natural wear, corrosion, and age-related issues. Material integrity is essential in determining a system’s ability to withstand ongoing use and environmental exposure. Key factors include equipment aging, structural degradation, wear and tear, and corrosion. Operational conditions are another key factor influencing the performance variability of technological functions. These include external conditions, such as environmental factors, that impact system operation. Stable environments offer predictability, while harsh conditions such as extreme temperatures or high humidity pose challenges that can compromise equipment functionality and increase variability. Table 4 presents the nodes, their respective states, and detailed descriptions.

Table 4: The identified contributing factors to the performance variability of technological functions.

Categories	Nodes	States	Descriptions	Reference
Safety-oriented factors	Maintenance strategy	Preventive-oriented, balanced, corrective-oriented	A preventive-oriented maintenance strategy emphasizes proactive measures to prevent potential failures, significantly boosting reliability but at a higher cost. In contrast, a corrective-oriented strategy addresses failures only after they occur. A balanced strategy combines both approaches, optimizing reliability while distributing the budget more evenly.	*(West et al., 2024)*
	Maintenance quality	Optimal, acceptable, poor	Maintenance quality evaluates the thoroughness and technical precision of maintenance tasks. Optimal maintenance quality reflects skilled execution, accuracy, attention to detail, and adherence to best practices and standards, while poor quality indicates a lack of these attributes.	*(Lu and Zhou, 2019)*
	Inspection practice	Intensive, moderate, sporadic	Sporadic or inadequate inspections raise the risk of undetected degradation, whereas an intensive inspection regimen enhances the detection of potential degradation.	*(Ferreira et al., 2009)*
	Maintenance effectiveness	High, moderate, low	Maintenance effectiveness refers to how successfully maintenance activities prevent or mitigate failures and ensure reliable operation of system components. It encompasses the impact of maintenance strategies, inspection frequency, and the quality of maintenance activities on equipment performance.	*(Costa and Cavalcante, 2022)*
	Reliability	High, moderate, low	Reliability indicates the system's likelihood to perform its function without failure, under a specified condition, and over a specified period of time.	*(Birolini, 2017)*
	Redundancy	Adequate, inadequate	Redundancy adds a layer of resilience; adequate redundancy reduces the likelihood of high variability in performance.	*(Peiravi et al., 2022)*
	MTTR	Short, long	Mean Time To Repair affects downtime; longer repair times increase the risk of performance interruptions.	*(Birolini, 2017)*
	Availability	High, moderate, low	Availability measures how often the system can perform its intended function, impacted by reliability, redundancy, and MTTR.	*(Birolini, 2017)*

Material integrity factors	Equipment aging	New, old	Equipment aging is the gradual decline in performance and reliability due to the natural lifecycle of components. with older equipment, it is more likely to exhibit variability in performance due to accumulated wear, reduced flexibility, and potentially outdated technology.	*(Clarotti et al., 2004)*
	Structural degradation	Low, moderate, high	Structural degradation captures the overall deterioration of components or subsystems due to a combination of internal stresses, environmental conditions, and aging. High levels of structural degradation pose significant risks to the system, leading to more frequent breakdowns, reduced load-bearing capacities, and increased variability in performance.	*(Wu et al., 2024)*
	Wear and tear condition	Minimal, moderate, severe	Mechanical wear and tear describe the progressive degradation of parts caused by continuous usage and friction over time. it affects performance and longevity, with severe wear leading to higher failure rates.	*(Moulahi and Zdiri, 2025)*
	Corrosion	Low, moderate, high	Corrosion impacts the integrity of materials, particularly metals and surfaces exposed to harsh environments. High corrosion rates significantly compromise structural strength, increase the likelihood of unexpected failures, and lead to reduced performance reliability.	*(Melchers, 2005)*
Operational context factors	Environmental conditions	Stable, variable, harsh	A stable environment features predictable and consistent conditions, with minimal fluctuations in factors like temperature, humidity, and air quality. In contrast, a harsh environment is marked by extreme or persistent stressors such as high temperatures, corrosive substances, heavy vibrations, high humidity, or dust. A variable environment exhibits moderate fluctuations in external conditions.	*(Duan et al., 2023)*

3.2.3. Human functions

Human functions, within the framework of the Safety-II concept, play a pivotal role as they offer the most flexibility to adapt to variability and mitigate its adverse effects on the overall system. Consequently, modelling human performance becomes a crucial component of systemic risk analysis in a CSTS. Numerous Human Reliability Analysis (HRA) methods have been developed in the literature to address this challenge *(Patriarca et al., 2020)*. Among these, the Cognitive Reliability and Error Analysis Method (CREAM) stands out as the most suitable for this study due to the following reasons:

I. Systemic perspective: CREAM is aligned with modern systemic approaches, such as the Safety-II concept, by examining both successful and erroneous human actions, rather than focusing solely on failures *(Hollnagel, 1998)*.
II. Versatility and applicability: CREAM is adaptable across various industries and contexts. It evaluates the interactions between human, technological, and organizational factors, making it an ideal tool for analysing CSTS *(Pei et al., 2024)*.
III. Context-sensitive analysis: The methodology integrates the impact of context on human performance using Common Performance Conditions (CPCs), enabling a detailed and situational understanding of reliability *(Sun et al., 2012)*.
IV. Focus on cognitive processes: Unlike traditional HRA methods that emphasize physical tasks, CREAM prioritizes cognitive functions such as decision-making and problem-solving, which are crucial in today's complex systems *(Huang et al., 2025)*.
V. Output compatibility with performance variability: CREAM's output, represented by Contextual Control Modes (CCMs), aligns seamlessly with the four types of performance variability outlined in this study: ST, LV, MV, and HV *(Kannally et al., 2025; Shi et al., 2023)*.

Building on the aforementioned reasons and drawing inspiration from the work of Yang et al. (Yang et al., 2013), this study applies a modified CREAM methodology to assess the SVI of human functions through a five-step sequence.

In the first step, various CPCs are described, along with their potential states and how they influence human performance reliability. The original CPCs are divided into nine categories (Hollnagel, 1998). In this study, a minor modification is introduced which replaces the "time of day" CPC with "circadian rhythm and stress." This change highlights the significant impact that sleep deprivation or misalignment with natural circadian cycles can have on performance. Unlike the "time of day" classification, which is based on fixed time intervals like day and night, the circadian rhythm considers biological phases that influence cognitive performance and alertness. This approach provides a more accurate reflection of how these factors affect human performance reliability. Table 5 presents the CPCs along with the associated details *(Xi et al., 2017; Zhou et al., 2018)*.

Table 5: CPCs description, their states, and effects.

CPC	CPC states	Effects
1) Training and competence (TAC)	Inadequate ($S_{1,1}$)	Negative
	Adequate with limited experience ($S_{1,2}$)	Neutral
	Adequate with high experience ($S_{1,3}$)	Positive
2) Human-machine interface and operational support (HMI)	Inappropriate ($S_{2,1}$)	Negative
	Tolerable ($S_{2,2}$)	Neutral
	Adequate ($S_{2,3}$)	Neutral
	Supportive ($S_{2,4}$)	Positive
3) Availability of procedures and plans (APP)	Inappropriate ($S_{3,1}$)	Negative
	Acceptable ($S_{3,2}$)	Neutral
	Appropriate ($S_{3,3}$)	Positive
4) Conditions of working (COW)	Incompatible ($S_{4,1}$)	Negative
	Compatible ($S_{4,2}$)	Neutral
	Advantageous ($S_{4,3}$)	Positive
5) Number of goals and conflict resolution (NGC)	More than actual capacity ($S_{5,1}$)	Negative
	Matching current capacity ($S_{5,2}$)	Neutral
	Fewer than actual capacity ($S_{5,3}$)	Positive

6) Available time and time pressure (ATT)	Continuously inadequate ($S_{6,1}$)	Negative
	Temporarily inadequate ($S_{6,2}$)	Neutral
	Adequate ($S_{6,3}$)	Positive
7) Circadian rhythm and stress (CRS)	High ($S_{7,1}$)	Negative
	Moderate ($S_{7,2}$)	Neutral
	Low ($S_{7,3}$)	Positive
8) Team collaboration quality (TCQ)	Deficient ($S_{8,1}$)	Negative
	Inefficient ($S_{8,2}$)	Neutral
	Efficient ($S_{8,3}$)	Neutral
	Very efficient ($S_{8,4}$)	Positive
9) Quality and support of the organization (QSO)	Deficient ($S_{9,1}$)	Negative
	Inefficient ($S_{9,2}$)	Negative
	Efficient ($S_{9,3}$)	Neutral
	Very efficient ($S_{9,4}$)	Positive

In step 2, the relationships between CPCs and CCMs are established by defining specific rules. These rules determine how various combinations of CPCs, along with their corresponding effects, influence the assigned values of the CCMs. The CCM, which represents the context of human cognition and action, is characterized by four distinct states: “strategic,” “tactical,” “opportunistic,” and “scrambled.” These relationships are formulated as if-then rules, where the "if" component specifies different CPC combinations and their effects, and the "then" component maps these combinations to the appropriate CCM characteristics.

In step 3, belief degrees are assigned to the consequences, or the "THEN" components of the rules, to account for uncertainty and ensure that minor variations in the "IF" components are accurately reflected in the "THEN" outcomes. To achieve this, a systematic approach is employed to determine the belief degrees by leveraging the basic control mode diagram of CREAM and a weighting system. The AHP is used to calculate the relative weights of all CPCs based on their importance. Subsequently, the conditional belief degrees, denoted as β^+ and β^-, are derived using the diagram shown in Figure 2. These degrees correspond to the positive or negative effects of various CPC states (Konstandinidou et al., 2006). To clarify the approach, an illustrative example is presented in Appendix A.

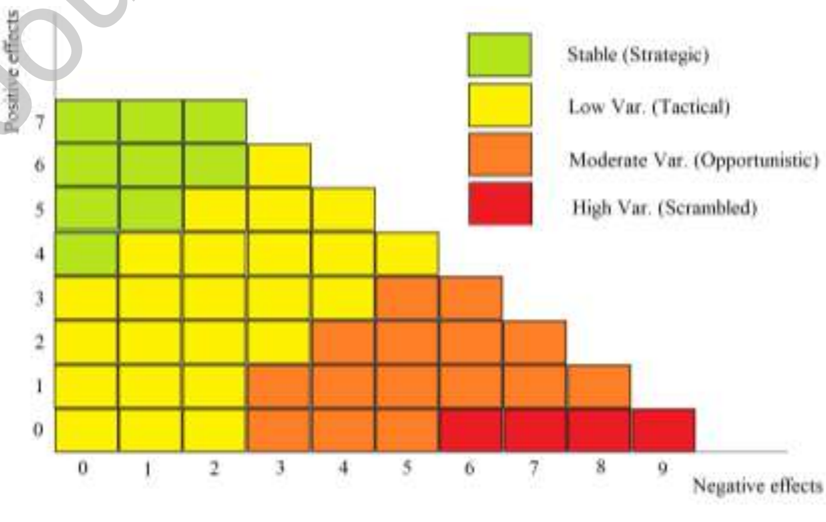

Figure 2: Basic Diagram of CREAM for different CCMs.

Step 4 involves constructing a BN to model the dependencies between CPCs. While CPCs share similarities with Performance Shaping Factors (PSFs) in other HRA methods, they are not the same. Their interdependencies are based on their influence on human performance reliability. Furthermore, CPCs may be calibrated based on the states of other CPCs. For instance, if a CPC initially exhibits a neutral effect but depends on other CPCs, its primary effect may shift toward either a positive or negative influence depending on the states of the CPCs it relies on. Table 6 illustrates the dependencies among various CPCs. The CPCs listed in the left-hand column are influenced by those defined in the top row. For instance, examining the third column reveals that “COW,” “NGC,” and “ATT” depend on “HMI.” This implies that if the human-machine interface and operational support improve, working conditions and the availability of time are expected to improve, as indicated by the letter “P,” representing a positive influence. Conversely, the number of goals and conflict resolution tasks required of the operator are expected to decrease, as denoted by the letter “N,” signifying a negative influence. The remaining cells in the table, marked with “-,” indicate no dependencies between the respective CPCs.

Table 6: Dependencies among CPCs.

	TAC	HMI	APP	COW	NGC	ATT	CRS	TCQ	QSO
TAC	-	-	-	-	-	-	-	-	P
HMI	-	-	-	-	-	-	-	-	P
APP	-	-	-	-	-	-	-	-	P
COW	P	P	-	-	-	P	P	-	P
NGC	-	N	N	N	-	-	-	-	-
ATT	-	P	P	P	N	-	P	P	-
CRS	-	-	-	-	-	-	-	-	-
TCQ	P	-	-	-	-	-	-	-	P
QSO	-	-	-	-	-	-	-	-	-

Considering these dependencies and the dynamic adjustability of CPCs based on the status of other related CPCs, a BN is well-suited for modelling these variabilities and interactive relationships. In this framework, the child node of the BN represents CCMs, which include the four defined states: strategic, tactical, opportunistic, and scrambled. These states align closely with the four performance variability levels commonly applied in both technological and organizational functions: ST, LV, MV, and HV, respectively. The parent nodes, representing the CPCs with their multiple states, are outlined in Table 5. To account for the dependencies shown in Table 6, four additional nodes, referred to as calibrated nodes, were introduced. These nodes capture the interactive relationships among CPCs and reflect their updated status based on changes in related CPCs. The four calibrated nodes are labelled as “calibrated COW”, “calibrated TCQ”, “calibrated NGC”, and “calibrated ATT”.

In the final step, the BN inference and validation process is carried out. This process includes determining the posterior probabilities of the target variables in the network and verifying the outcomes to confirm the precision and dependability of the suggested approach. First, observations are analysed to derive the prior probabilities for each CPC in terms of numerical variables that correspond to CPC states and their effects. Next, during the inference process, belief degrees are converted into rules, which serve as the conditional probabilities for the constructed BN. Using these transformed rules and the prior probabilities, the marginal probabilities of the leaf node states are then computed accordingly.

3.2.4. **Prior probabilities extraction**

Due to the complexity of CSTS and the diverse nature of their elements, various data sources with different origins are required to inform the developed models. For technical functions, several data types are particularly useful. Measurements from equipment sensors, operational conditions, and processes provide valuable empirical data. Operational logs detailing equipment performance and failures are essential, as are records of preventive and corrective maintenance activities, which help evaluate maintenance effectiveness. Additionally, manufacturer specifications, including reliability data such as Mean Time to Failure (MTTF), Mean Time to Repair (MTTR), and other relevant metrics, are integral to reliability assessment.

When it comes to organizational functions, obtaining realistic data can be challenging. Managers are often reluctant to critique their management practices, organizational structure, or operational efficiency due to concerns about reputation and prestige *(Liu, 2021; LÜScher and Lewis, 2008)*. Nevertheless, for the organizational functions, valuable information can be gathered from various sources, including compliance and incident data from internal audits, human resource databases (e.g., staff turnover rates, training schedules, and role-specific records), and regulatory databases containing compliance reports or industry-level performance benchmarks. Additionally, input from independent expert elicitation can be incorporated for several nodes of the developed BN. In relation to the structure of organizational performance, organizational resources, and external factors, the data are primarily obtained from documented evidence and available empirical sources. However, obtaining objective data on organizational management and culture remains inherently challenging, as such aspects are often subjective and difficult to quantify even across other industrial sectors. For instance, safety culture is a latent and intangible construct that cannot be measured directly. It is typically assessed through a combination of subjective (survey-based) and objective (performance-based) indicators. Although no purely objective measure of safety culture exists, triangulating multiple data sources, such as surveys, audits, and performance indicators, enhances validity and reduces bias. Accordingly, for the organizational functions, both empirical and subjective data sources are employed to capture the multifaceted nature of organizational performance.

Assessing human performance variability requires the use of expert judgment, as databases in this area are often insufficient to meet expectations. To this end, the Dempster-Shafer evidence theory (DSET) is employed for several purposes:

a) Systematically combining diverse expert opinions to produce a unified final judgment.
b) Accounting for both epistemic and aleatory uncertainties, thanks to its unique features, such as representing and propagating degrees of belief.
c) Providing a structured framework for reasoning under uncertainty, allowing for the integration of incomplete or conflicting evidence.

This approach enhances the reliability of expert-based assessments by managing variability and uncertainty in a more systematic and robust manner. DSET is frequently characterized as an advanced form of probability theory or an expanded interpretation of Bayesian inference. It has been widely used to extract subjective expert judgments and resolve disparities between differing viewpoints to produce an aggregated output. In this context, DSET is referred to as a theory of evidence because it focuses on the weight of evidence. Before combining information, the foundational principles of DSET must be introduced. A comprehensive explanation of DSET can be found in the literature (Gros, 1997; Tang et al., 2023), while a brief introduction is provided in Appendix B.

3.3. Quantitative analysis of system performance variability

Once the FRAM model is developed and the internal variabilities across all function categories are obtained, the next step is to map the FRAM model into a BN to conduct a quantitative analysis of system performance variability. As previously discussed, various types of variabilities are integral to an FRAM model, including SVI and UVI. Aggregating these variabilities across different functions is essential to gain a comprehensive understanding of performance variability within a CSTS. This aggregation represents the unified interactions between functions that are interconnected in a sequential manner within the FRAM model.

3.3.1. FRAM and BN integration

The process begins by converting various aspects of a function into discrete probability distributions, categorized into states such as ST, LV, MV, and HV. This approach enhances the representation of functional variability and serves as a common framework, simplifying the interpretation of interactions between functions (Patriarca et al., 2017). Furthermore, the internal variability identified for each function can be regarded as an additional dimension, reflecting the influence of the operational environment and current performance conditions during the function's execution (Slim and Nadeau, 2020). The mapping process begins with the output from background functions, establishing the initial performance variability distribution for downstream functions. This variability can be determined either through empirical data, if available, or expert elicitation when data is limited. To represent this as discrete probability distributions, the frequency of event occurrences may be used when empirical data is applied. For each function, all available and defined aspects are set as parent nodes in the BN model, with the output serving as the child node. This configuration enables a quantitative calculation of the interactions among different aspects of each function, resulting in an integrated performance variability distribution with consistent state definitions. Figure 3 demonstrates a simplified mechanism for mapping the FRAM model onto a BN, providing clearer insight into the process.

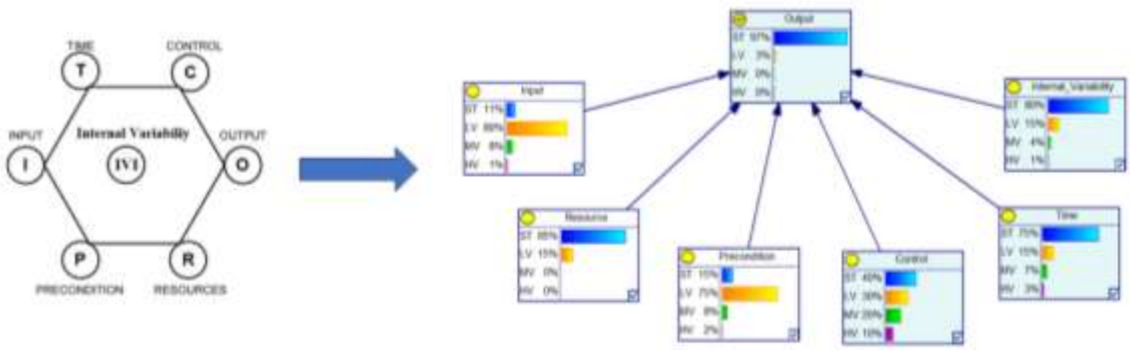

Figure 3: The simplified process of mapping FRAM into a BN model.

A key advantage of BN is its flexibility in integrating a variety of nodes with multiple states, accommodating both discrete and continuous forms. Given this flexibility and recognizing performance variability across four distinct states defined by a discrete probability distribution, as well as the independence of different functional aspects and their separate impacts on the output, the CPT can be calculated using canonical probabilistic models like noisy OR, noisy MAX, noisy MIN, noisy AND, and noisy Adder gates (Diez and Druzdzel, 2006). The Noisy-OR model, introduced by Pearl (1988), initially addressed probabilistic dependencies among binary variables. (Henrion, 1989) extended this concept, adapting the model to include binary leaky Noisy-OR gates, which account for additional uncertainty in influence pathways. Further developments came when *(Díez, 1993)* and Srinivas (1993) independently proposed generalizations of the model to accommodate multi-valued variables, leading to the creation of multi-valued Noisy-OR gates. These foundational works eventually paved the way for the Noisy-MAX model, which expanded the framework to capture more complex probabilistic relationships across diverse variable states. In this study, the complexity of the problem, characterized by multi-state parent nodes, a multi-state child node, and the independent influence of each parent on the child makes the Noisy-MAX technique particularly suitable. This approach not only streamlines the construction of the CPT but also effectively captures the non-linear relationships between parent and child nodes, enabling a more accurate representation of these dependencies *(Cantelmi et al., 2025; Xie et al., 2024; Xue et al., 2025).*

3.3.2. Noisy-MAX structure-based BN modelling

Using the Noisy-MAX technique, the conditional probability between a child node C and its parent node R can be represented by incorporating a set of n auxiliary variables $\{A_{1,}\ldots,A_n\}$(Diez and Druzdzel, 2006). As illustrated in Figure 4, this formulation allows the conditional probability to be expressed as:

$$P(C/R) = \sum_{A} P(C/A).P(A/R) \quad (1)$$

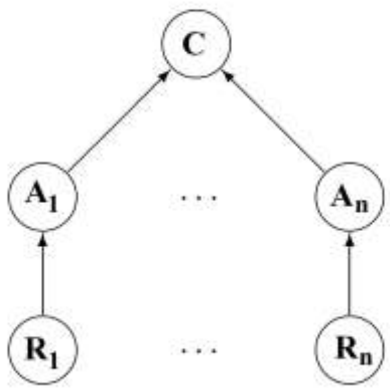

Figure 4: Simplified BN structure for noisy-MAX model derivation.

Note that the variables A_i are purely auxiliary elements used to facilitate equation derivation and are not part of the actual model. Given the graph in Figure 4, there are no interactions between the causal mechanisms through which R_i influences the value of C. In the graph, this property is represented by the absence of connections $R_i \rightarrow A_j$ and $A_i \rightarrow A_j$ for all $i \neq j$, indicating that:

$$P(A/R) = \prod_i P(A_i/R_i) \quad (2)$$

With this, combined with Equations 1 and 2, results in:

$$P(C/R) = \sum_{A/f(A)=C} \prod_i P(A_i/R_i) \quad (3)$$

In this context, each A_i signifies the contribution of R_i to the value of C. The combined outcome generated by each R_i is represented as $C=f_{MAX(A)}$. Consequently, C and A_i variables must operate within the same domain. Each A_i reflects the impact of R_i elevating C to a particular level, and the actual value of C is determined as the maximum among the A_i values.

Now, to establish the CPT for the Noisy-MAX model, we must calculate $P(C=c/A)$ for every possible value c and each configuration of R. This is achieved by applying Equation 3 and recognizing that $f_{MAX(A)=}max(A_1,\ldots,A_n)$. This function implies that $f_{MAX(A)} \leq y$, if and only if $A_i \leq C$ for each i. Therefore, we have:

$$P(C \leq c/R) = \sum_{A/f_{MAX(A)} \leq c} \prod_i P(A_i/R_i) = \sum_{A_1 \leq c} \ldots \sum_{A_n \leq c} \prod_i P(A_i/R_i) = \prod_i \left(\sum_{A_i \leq c} P(A_i/R_i) \right) \quad (4)$$

With consideration of accumulative parameters, the values of the CPT can be obtained as follows:

$$P(c/R) = \begin{cases} P(C \leq c/R) - P(C \leq c-1/R) & for\ c \neq c_{min} \\ P(C \leq c/R) & for\ c = c_{min} \end{cases} \quad (5)$$

After calculating the CPTs for all BNs related to each function, using prior probabilities derived from either empirical data or expert input, the complete set of BN models is analysed to generate the final output for the last function. This analysis enables us to assess the performance variability of each function independently, as well as to evaluate its impact on downstream functions.

3.3.3. FRAM interpretation process

The ultimate goal of FRAM modelling is to understand how disruptions or variations in upstream functions influence the performance variability of connected functions. In essence, it examines how resonance, whether positive or negative, affects the variability in performance across downstream functions. This approach provides a detailed view of how any disruption in a CSTS can propagate, helping us understand how changes in one part of the system influence the entire system's behaviour. To this end, a 2-D criticality matrix is proposed to support the decision-making process *(Kaya et al., 2021; Patriarca et al., 2018; Zarei et al., 2022)*. The matrix dimensions represent probability and consequence. For the probability dimension, the mean value of performance variability serves as a numerical representation of the average variability a function experiences. This considers the likelihood of being in one of four states: HV, MV, LV, or ST, multiplied by the assigned scores of 4, 3, 2, and 1, respectively. These scores reflect the significance of each state in terms of safety impact. HV is given the highest score (4) to represent substantial disruption; MV receives a moderate score (3) for moderate variability; while LV and ST are assigned lower scores (2 and 1) to indicate minimal variability or stability. For the consequence dimension, three categories are defined: critical (indicating severe consequences), moderate (manageable consequences requiring attention), and minor (minimal or tolerable consequences). Functions are classified into these categories based on their significance to both safety and operational performance. The magnitude of consequences is highly dependent on the specific domain under study and the function's role in the system's operation and safety. This classification can be determined using expert judgment or established criteria. Figure 5 illustrates the proposed criticality matrix, which categorizes functions into three levels of criticality based on their variability and consequence severity.

Variability level	Severity level: Minor	Moderate	Critical
HV	B	C	C
MV	B	B	C
LV	A	B	B
ST	A	A	B

Figure 5: The proposed criticality matrix.

Level C, located in the top-right quadrant, represents high variability and critical consequences. Functions in this category are prime candidates for triggering negative resonance, as their high variability combined with critical consequences makes them likely to interact unpredictably with downstream functions, potentially amplifying risks across the system. Level B, which includes functions with moderate variability, highlights that these functions can also contribute to negative resonance. This occurs particularly when their variability interacts with other moderately variable or interconnected functions, creating conditions where risks propagate through the system. Such interactions are especially critical when these functions are linked to others with similar variability characteristics. Level A encompasses functions that are relatively stable or exhibit low variability. These functions can play a stabilizing role within the system and be strategically leveraged to design interventions that dampen variability and mitigate risks. By strengthening the interactions of these stable functions, they can counteract the effects of high variability in connected functions. This criticality matrix provides a systematic tool to prioritize functions for intervention based on their role in system dynamics. It facilitates the detection and mitigation of resonances in the FRAM model by anticipating how function interactions might lead to either risk amplification (negative resonance) or system stabilization (damping resonance).

3.4. Verification and validation process

Verification and validation process are the essential component of any methodological approach, ensuring that developed models are reliable, robust, and sensible. They also builds confidence in the accuracy of the findings and results. In this study, various techniques and numerous models have been employed to address the complexity of CSTS, making comprehensive benchmarking challenging. To address this, we adopted a modular approach using a range of techniques, allowing us to validate and benchmark different models independently. Validation of the HTA and FRAM components, as qualitative analysis methods, primarily depends on the knowledge and proficiency of the analysts conducting the evaluation. Additionally, the results and findings from these models are compared and benchmarked against outcomes from similar studies.

For validating the developed BN models, sensitivity analysis, regarded as one of the most practical validation methods, is applied. This analysis involves two approaches. The first approach confirms the model's robustness by verifying that small adjustments in the prior probabilities of parent nodes reliably affect the probabilities of child nodes. This principle-based sensitivity analysis ensures that the model responds predictably to changes in inputs, enhancing its reliability and accuracy. To achieve this, the analysis follows these principles (Jones et al., 2010):

- Principle 1: Minor adjustments in the prior probabilities of the parent node should lead to proportionate changes, either increases or decreases, in the posterior probability distribution of the child node.
- Principle 2: The combined influence of changes in the probabilities of all evidence variables should be equal to or greater than the influence produced by modifying any individual subset of that evidence.

In the second approach, the analysis focuses on how changes in probability parameters influence the BN's output. This is done by calculating the derivatives of the posterior probability distributions, which helps reveal the sensitivity of the model's target nodes (such as performance variability) to adjustments in various numerical parameters. This derivative-based analysis measures the rate at which each target node's probability shifts as a reaction to small modifications in the parent nodes' prior probability values. By examining these derivatives, the parameters that most strongly influence the network's outcomes can be identified. When certain variables show high sensitivity to parameter changes, it indicates that the model depends significantly on those specific inputs. Recognizing these key parameters allows for prioritizing data that may require more precise estimates or rigorous validation, as they play a crucial role in determining the model's predictions. It is noted that for ease of reference, all symbols and mathematical notations used in the proposed methodology are summarised in the Table I in Appendix C.

4. Results, discussion, and implications

Seaports are widely regarded as a CSTS that are highly interconnected and interdependent, making them vulnerable to a diverse range of risks. Given that reliable and efficient seaport operations are essential for the maritime transportation sector, any disruptions or fluctuations in their performance can significantly impact national safety, security, economic stability, and public health *(Mohsendokht et al., 2025)*. This underscores the critical need for focused attention from risk analysts to develop robust approaches to address these challenges. This section applies the proposed methodology to a typical seaport, illustrating both its practicality and potential impact.

4.1. FRAM model development

To identify the key functions for FRAM development, an initial HTA is conducted to represent the workflow of activities typically performed in a seaport. The hierarchical structure of the HTA provides a comprehensive understanding of the workflow and facilitates a detailed analysis of specific tasks along with their prerequisite requirements. It is important to note that seaport operations involve a vast array of tasks and the collaboration of numerous teams and crews (Carlo et al., 2015; Haas, 2016). To maintain simplicity and align with the scope of a journal paper, a streamlined version of the HTA focusing on the most critical activities is produced. The HTA was developed by synthesizing insights from an extensive review of the seaport operations literature, the collective research contributions of the author team, and subsequent verification and approval by a panel of experts whose profiles are provided in Table II in Appendix C. It should be noted that this study focuses solely on operations occurring between the quay area and the yard within the seaport. Figure 6 presents this simplified HTA, which serves as the foundation for the FRAM model.

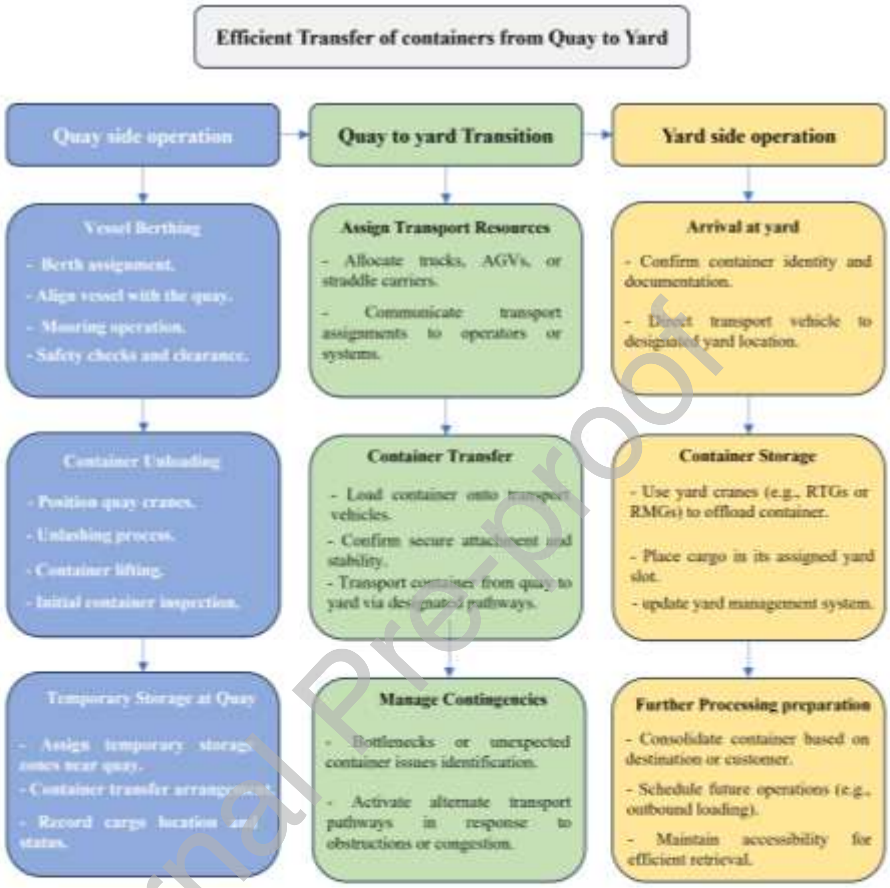

Figure 6: HTA for seaport activities.

Based on the HTA results, nine foreground functions, referred to as main functions, and four background functions have been selected for the FRAM development. The background functions define the boundaries of the analysis, providing fixed outputs that feed into and support the operation of the main functions. Table 7 outlines the functions, their characteristics, and the connections between them, while Table 8 details the various aspects of each function.

It is noted that system complexity increases rapidly with size, rendering manual modelling increasingly challenging for large infrastructures such as seaports. To address this issue, the FRAM model is organized into interacting modules, such as quayside operations, yard operations, and intermodal transfer sections, structured across hierarchical levels, namely Macro, Meso, and Micro, as illustrated in Figure 7.

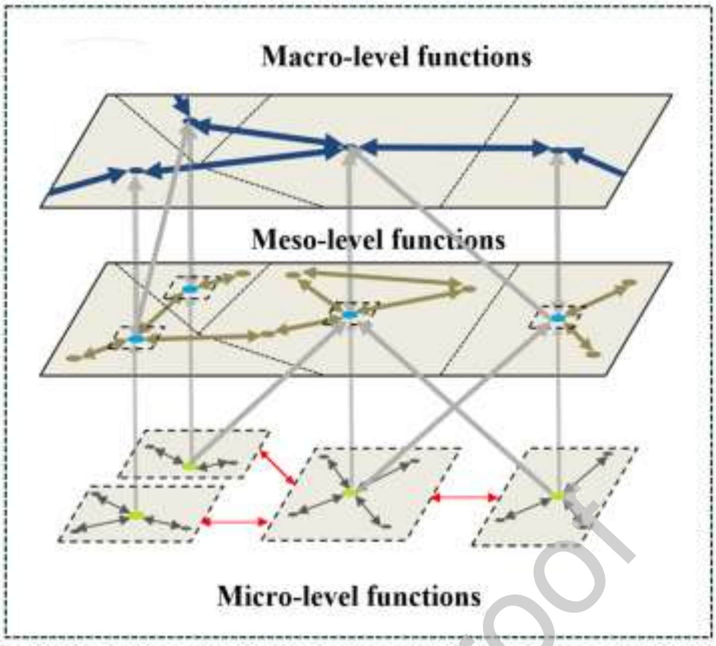

Figure 7: The interacting hierarchical levels in seaport operations.

At the Macro level, the focus is on the seaport as an integrated system, capturing the high-level interactions between major operational sections and strategic processes, such as overall cargo throughput, terminal coordination, and resource allocation. The Meso level examines intermediate-scale processes within individual modules. For example, within the quayside module, functions such as quay crane operations, vessel berthing, and container handling are considered, along with their interconnections and resource flows. At the Micro level, the model focuses on detailed, function-specific activities, including the interactions between individual equipment, human operators, and tasks. For instance, a micro-level analysis of a quay side may include the container unlashing process, operator actions, and resource availability.

Each module is analysed largely independently, with only a limited set of interface variables connecting it to other modules. This divide-and-compose strategy contains complexity locally, prevents combinatorial growth as system size increases, and ensures that both high-level coordination and fine-grained operational variability are effectively represented.

All the identified functions and their interconnections are synthesized and visualized using the FRAM Model Visualization (FMV) tool (Hollnagel et al., 2023), as shown in Figure 8.

Table 7: Function description, characterization, and links.

Function	Description	Type	Links
F1	Berth assignment and confirmation	Organizational	F1(O)→F2(I)
F2	Initial Safety and Security Checks	Human	F2(O)→F3(I)
F3	Unlashing of Containers	Human	F3(O)→F4(I)
F4	Cargo Unloading Preparation	Organizational	F4(O)→F5(I), F6(I)

F5	Quay crane operation	Technological	F5(O)→F7(I)
F6	Quay crane operator	Human	F6(O)→F5(C)
F7	Cargo Transport to Yard Storage	Technological	F7(O)→F8(I), F9(I)
F8	Yard crane operator	Human	F8(O)→F9(C)
F9	Yard crane operation	Technological	-
BG1	Vessel securely moored	Background function	BG1(O)→F2(P)
BG2	Port operations management	Background function	BG2(O)→F1(C), F2(C), F3(C), F7(C), F8(C), F9(C)
BG3	Berth assignment information	Background function	BG3(O)→F1(I)
BG4	Resource management	Background function	BG4(O)→ F1(R), F2(R), F3(R), F4(R), F5(R), F6(R), F7(R), F8(R), F9(R)

Table 8: Functions aspects descriptions.

Function	Output	Input	Pre-condition	Control	Resource
F1	Confirmation of berth assignment	Berth assignment information	-	Port authority protocols	Communication systems, Port staff
F2	Safety and security status report	Confirmation of berth assignment	Vessel securely moored	Port security regulations	Safety and security equipment, Personnel (security officers)
F3	Unlashed containers ready for unloading	Safety and security status report	-	Unlashing protocols, Supervisor instructions	Unlashing tools, Personnel (dock workers)
F4	Instructions for crane operators, Updated cargo status	Unlashed containers	-	Port operations management, Communication from the vessel	Communication systems
F5	Cargo unloaded to dock	Updated cargo status	-	Crane operator's commands, Crane control system	Crane and fuel/power supply, Operator
F6	Crane operator's commands	Instructions for crane operators	-	-	Communication systems
F7	Cargo delivered to yard storage	Cargo unloaded to dock	-	AGV control management system	Transport vehicles (e.g., AGVs, trucks), Drivers and handlers
F8	Crane operator's commands	Cargo delivered to yard storage	Clear storage allocation instructions, Safety checks completed	-	Communication systems
F9	Cargo properly placed in designated storage areas	Cargo delivered to yard storage	-	Yard management system, Operator commands,	Crane and fuel/power supply, Operator

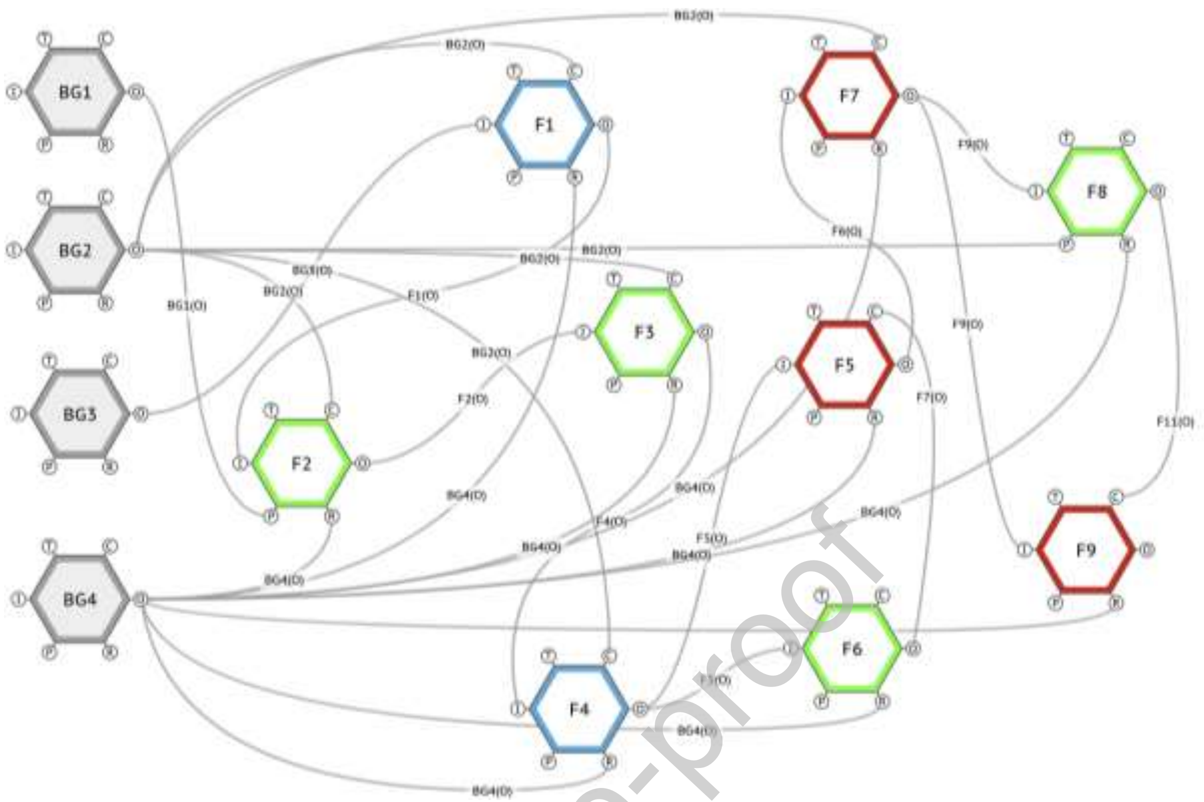

Figure 8: The FRAM model of typical activities conducted in a seaport.

4.2. SVI assessment for key functions

4.2.1. Organizational functions

In the context of seaport organizational functions, numerous entities are involved, with complex interactions among their components. To assess the performance of their internal variability, a BN for the organizational function is constructed, following the information and framework described in Section 3.2.1. As shown in Figure 9, the performance variability of an organizational function is influenced by five intermediate nodes: organizational culture, organizational management, organizational resources, organizational structure, and external factors. Each of these intermediate nodes is determined by its respective parent nodes. Achieving a stable condition with a high probability requires all intermediate nodes to be in their most favourable states. This includes having a highly efficient organizational structure, sufficient and well-allocated resources, optimal organizational management practices, a rich and supportive organizational culture, and minimal impact from external factors. On the other hand, highly variable organizational performance arises when the intermediate nodes are in their least favourable states. For instance, an inefficient structure, inadequate resources, poor management, a weak organizational culture, and significant external pressures collectively lead to increased variability in performance. This relationship underscores the importance of maintaining favourable conditions across all intermediate nodes to ensure organizational stability.

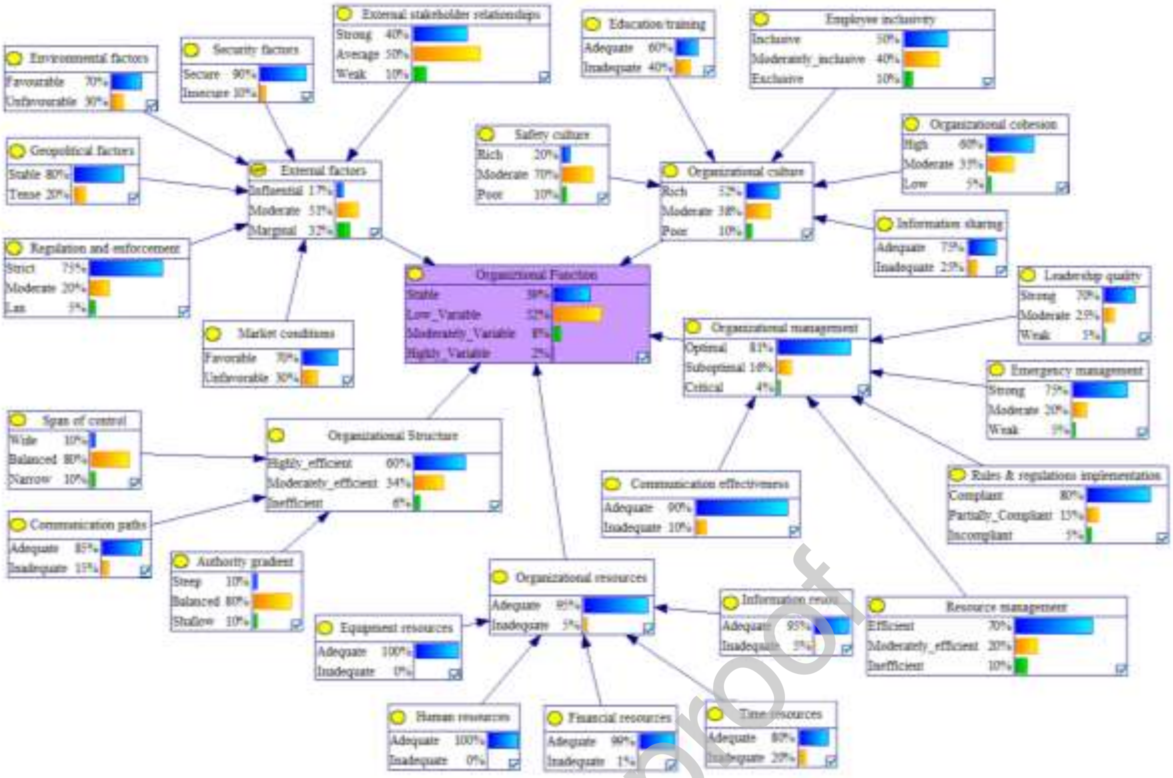

Figure 9: The BN model for SVI assessment of organizational functions.

4.2.2. Technological functions

In a seaport, various types of machinery, equipment, and their components contribute to the activities of technological functions. To evaluate their internal variability performance, the corresponding BN for each technological function is developed based on the information and structures outlined in Section 3.2.2. Due to space constraints in the journal paper format, only the BN for quay cranes is presented in Figure 10 to demonstrate the applicability of the proposed methodology. Quay cranes are widely regarded as the most important, valuable, costly, and complex components in a seaport. A seaport without them is often considered paralyzed, as they serve as the critical link between sea and land operations.

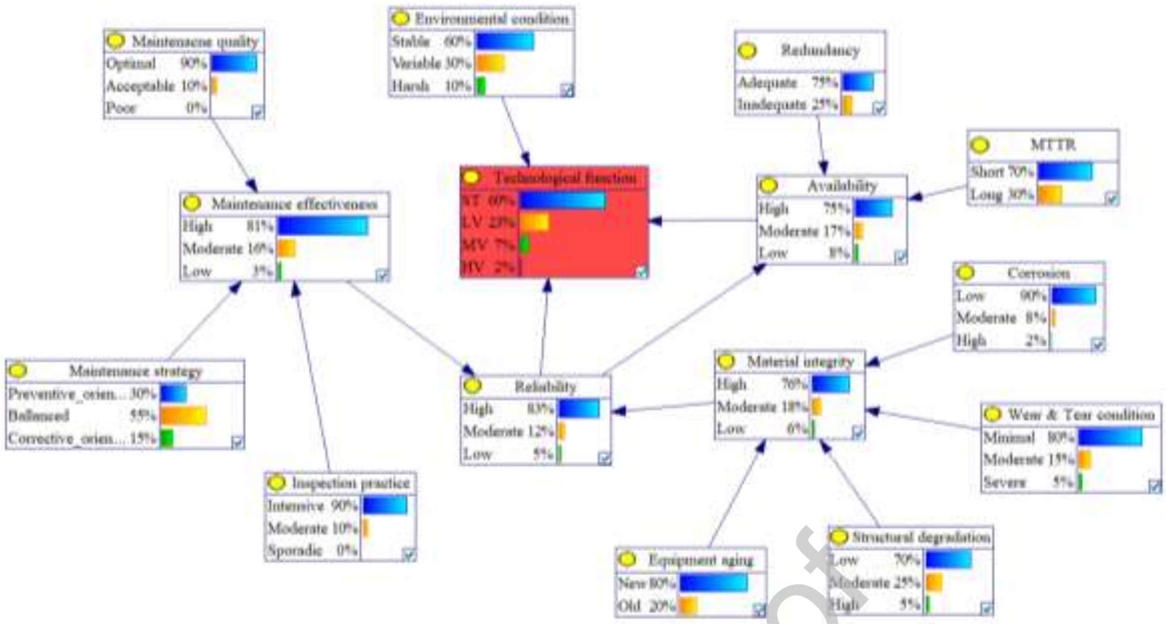

Figure 10: The BN model for SVI assessment of technological functions.

In the developed BN, maintenance effectiveness and material integrity are identified as the two key factors directly influencing equipment reliability. Greater levels of material integrity and maintenance effectiveness correspond to higher reliability. It is noteworthy that the operation of complex systems such as quay cranes often involves dynamic processes that impact their structure and the reliability of their components over time. Given the critical importance of ensuring both safety and operational effectiveness, a shift from a two-state to a multistate approach in reliability analysis is warranted. This approach facilitates a more accurate assessment of their dependability and operational effectiveness. It also helps identify critical reliability thresholds, where exceeding these limits may fail to ensure the required level of operational effectiveness (Kołowrocki and Soszyńska-Budny, 2011). Therefore, the reliability is categorized into three states: high, moderate, and low, defined according to the specific characteristics of the component in question. For quay cranes, high reliability corresponds to a reliability level between 95% and 100%, moderate reliability falls between 85% and 95%, and low reliability is defined as below 85% *(Deng, 2000; Jo and Kim, 2019)*. Availability is determined by three key factors: reliability, MTTR, and redundancy. Higher reliability and redundancy contribute to increased availability, while a shorter MTTR enhances availability by reducing equipment downtime. Technological performance variability depends on three factors: reliability, availability, and environmental conditions. The SVI is likely to remain stable with high probability if environmental conditions are stable and both reliability and availability are high. Other SVI states are assigned proportional values based on the probabilities of their parent states.

To illustrate the applicability of the methodology, prior probabilities were derived from historical records of the seaport under study, representing its current status. As depicted in Figure 10, the stable state of the technological function is assigned a probability of 68.8%, while the remaining probabilities are distributed as follows: 22.7% for the LV state, 6.7% for the MV state, and 1.8% for the HV state. These values reflect the system's realistic behaviour, highlighting the influence of various factors that create discrepancies between "work as imagined" and "work as done."

4.2.3. Human functions

To determine the SVI for human functions, the modified CREAM methodology outlined in Section 3.2.3 is employed. The process begins with developing the BN structure by identifying the main CPCs, their interdependencies, and incorporating calibrated CPCs. The leaf node in the network is represented as the CCM, which reflects human action status. The four well-known modes (strategic, tactical, opportunistic, and scrambled) are interpreted as ST, LV, MV, and HV, respectively. Figure 11 illustrates the resulting BN for human functions.

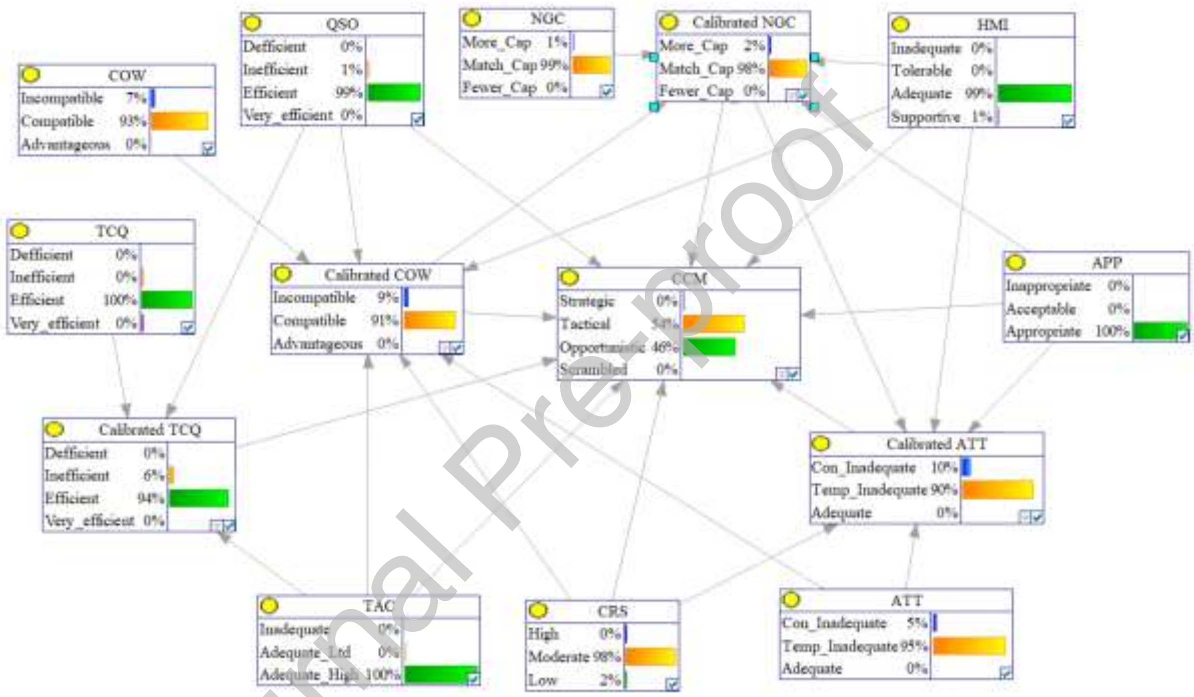

Figure 11: The BN model for SVI assessment of quay crane operator.

Next, the rules governing the BN are organized using a belief structure that accounts for all possible combinations of CPC states. These rules form the CPT for the developed BN. It is important to note that not all CPCs equally influence human performance variability. To address this, the AHP method is applied to determine appropriate weights for each CPC, tailored to the nature of tasks performed by humans in a seaport environment *(Yang et al., 2013)*. Table 9 presents a pairwise comparison matrix showing the weights for all nine CPCs. The consistency ratio, calculated as *6.7E-3*, confirms that the derived weights are logically consistent and represent a well-justified hierarchy of importance.

Table 9: Deriving CPC weights using AHP method.

CPC	TAC	HMI	APP	COW	NGC	ATT	CRS	TCQ	QSO	Weight
TAC	1.00	5.00	2.00	4.00	1.50	1.50	5.00	5.00	5.00	0.25
HMI	0.20	1.00	0.33	1.00	0.25	0.25	1.00	1.00	1.00	0.05
APP	0.50	3.00	1.00	2.00	0.67	0.67	3.00	3.00	3.00	0.13
COW	0.25	1.00	0.50	1.00	0.33	0.33	0.50	1.50	1.50	0.06

NGC	4.00	4.00	1.50	3.00	1.00	1.00	4.00	4.00	4.00	0.18
ATT	0.67	4.00	1.50	3.00	1.00	1.00	4.00	4.00	4.00	0.18
CRS	0.20	1.00	0.33	2.00	0.25	0.25	1.00	1.00	1.00	0.05
TCQ	0.20	1.00	0.33	0.67	0.25	0.25	1.00	1.00	1.00	0.05
QSO	0.20	1.00	0.33	0.67	0.25	0.25	1.00	1.00	1.00	0.05

After determining the weights, the rules with their corresponding belief degrees are established following the instructions in Section 3.2.3. To illustrate the process, Rule 23326 is used as an example. This rule is defined by the set ($S_{1,2}$, $S_{2,2}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,2}$), which corresponds to the effects (neutral, neutral, positive, positive, positive, positive, positive, positive, negative) based on the guidance in Table 5. Using Figure 2, the following values are subsequently calculated:

$$\beta^{+} = \{(0.75, ST), (0.25, LV), (0, MV), (0, HV)\}$$

$$\beta^{-} = \{(0.38, ST), (0.62, LV), (0, MV), (0, HV)\}$$

Using the corresponding weights of CPCs with positive effects from Table 9, their sum, W^{+} is calculated as *0.83 (=0.13+0.06+0.18+0.18+0.05+0.05)*. Conversely, W^{-} showing the weights of negatively influencing *CPCs*, is 0.05, since QSO is the only CPC with negative effect in the given set. Having obtained the normalized values of W^{+} and W^{-}, along with the corresponding β^{+} and β^{-} values using the evidential reasoning algorithm implemented in IDS software *(Xu and Yang, 2005)*, the final results for this combination of CPCs are determined as follows:

$$\beta = \{(0.745, ST), (0.255, LV), (0, MV), (0, HV)\}$$

In this manner, all the rules and their corresponding values are determined. Table 10 provides an example by showcasing nine of these rules, including the first three rows, three from the middle, and the last three rows.

Table 10: Rule-based CPT development for human function BN.

Rules	CPC combinations (IF part)	CCM belief degrees (THEN part)
1	$S_{1,1}$, $S_{2,1}$, $S_{3,1}$, $S_{4,1}$, $S_{5,1}$, $S_{6,1}$, $S_{7,1}$, $S_{8,1}$, $S_{9,1}$	{(0.000, ST),(0.000, LV),(0.000, MV),(1.000, HV)}
2	$S_{1,2}$, $S_{2,1}$, $S_{3,1}$, $S_{4,1}$, $S_{5,1}$, $S_{6,1}$, $S_{7,1}$, $S_{8,1}$, $S_{9,1}$	{(0.000, ST),(0.000, LV),(0.500, MV),(0.500, HV)}
3	$S_{1,3}$, $S_{2,1}$, $S_{3,1}$, $S_{4,1}$, $S_{5,1}$, $S_{6,1}$, $S_{7,1}$, $S_{8,1}$, $S_{9,1}$	{(0.000, ST),(0.030, LV),(0.561, MV),(0.409, HV)}
…	…	…
23326	$S_{1,2}$, $S_{2,2}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,2}$	{(0.745, ST),(0.255, LV),(0.000, MV),(0.000, HV)}
23327	$S_{1,2}$, $S_{2,2}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,3}$	{(0.750, ST),(0.250, LV),(0.000, MV),(0.000, HV)}
23328	$S_{1,2}$, $S_{2,2}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,4}$	{(1.000, ST),(0.000, LV),(0.000, MV),(0.000, HV)}
…	…	…
46654	$S_{1,3}$, $S_{2,4}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,2}$	{(0.998, ST),(0.002, LV),(0.000, MV),(0.000, HV)}
46655	$S_{1,3}$, $S_{2,4}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,3}$	{(1.000, ST),(0.000, LV),(0.000, MV),(0.000, HV)}
46656	$S_{1,3}$, $S_{2,4}$, $S_{3,3}$, $S_{4,3}$, $S_{5,3}$, $S_{6,3}$, $S_{7,3}$, $S_{8,4}$, $S_{9,4}$	{(1.000, ST),(0.000, LV),(0.000, MV),(0.000, HV)}

In the subsequent step, prior probabilities for various CPC states are determined based on expert judgment. Three seasoned experts were asked to assess the performance variability of quay crane operators during a typical yet busy day at a seaport, taking into account potential disruptive scenarios. The experts with extensive experience in seaport operations are interviewed to provide their probabilistic assessments, assigning values between 0% and 100% to different states. These individual judgments are then aggregated using DSET, yielding consolidated probabilities for each CPC state, as shown in Table III in Appendix C. These probability-based insights are incorporated into the BN as prior probabilities, while the rules and corresponding values in Table 10 serve as the CPT. To demonstrate the process, a sample calculation is provided as follows:

$$P(CCM) = \sum_{i=1,j=1}^{3} \sum_{i=2,j=1}^{4} \sum_{i=3,j=1}^{3} \sum_{i=4,j=1}^{3} \sum_{i=5,j=1}^{3} \sum_{i=6,j=1}^{3} \sum_{i=7,j=1}^{3} \sum_{i=8,j=1}^{4} \sum_{i=9,j=1}^{4} P(CCM/S_{i,j})P(S_{i,j}) \qquad (14)$$

Where *P(CCM)* represents the probability of performance variability in each of the four possible states, $P(CCM/S_{i,j})$ denotes the conditional probability of CCM given $S_{i,j}$, and $P(S_{i,j})$ signifies the probability of specific states of a given CPC. Depending on the aggregated probabilities and their potential impact on performance variability, proportional rules are extracted from Table 10. In this case, referring to the states of each CPC and their corresponding values, 576 rules are derived, representing various combinations of these states. By applying these rules and the values obtained from the aggregated expert judgments to the developed BN, the probabilities for the CCM states are calculated as follows: ST=0.005, LV=0.536, MV=0.459, and HV=0.000. These values suggest that, in the specified situation, the performance variability of the quay crane operator is more inclined toward low and moderate levels of variability.

4.3. FRAM to BN mapping results

After determining the SVI values for each function in the previous sections, the next step is to adopt a holistic perspective on the variabilities within the entire model. To achieve this, the output of each upstream function is integrated as the input or other related aspects for downstream functions. For illustration, Figure 12 highlights the output of Function 2. As shown, the input to Function 2 is derived from the output of Function 1, while other aspects of Function 2, along with its SVI, are represented as independent parent nodes in the developed BN. In this context, the noisy-max technique is utilized to calculate the CPT values for inter-functional relationships in the BN, as outlined in Section 3.3.2. The output performance variability of Function 2 results from the interaction of several factors: the input from Function 1, the SVI associated with Function 2, and the contributions of background functions BG1, BG2, and BG4, which serve as the precondition, control, and resources, respectively.

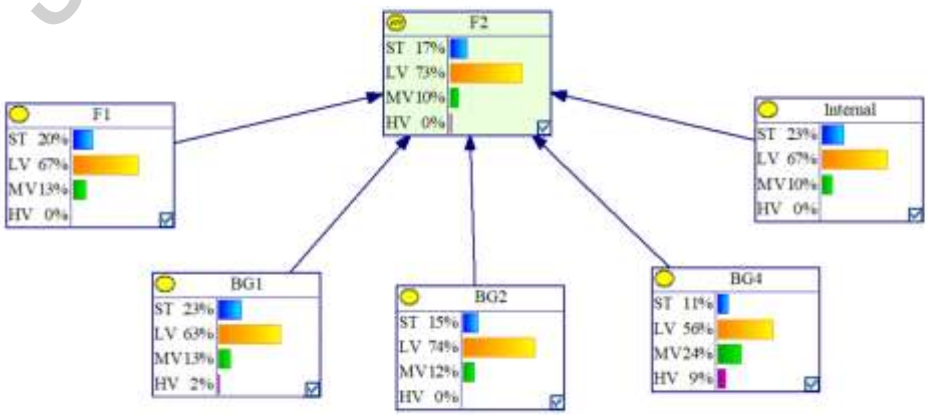

Figure 12: BN model for UVI calculation in function 2.

The values for background functions, which define the boundaries of the developed FRAM model, are derived using various methods discussed in this paper. For BG_1, representing the performance variability of mooring operations in a seaport, these values are determined from empirical data collected over several years. The ST state corresponds to all mooring operations that were conducted successfully and safely, adhering to the plan without any disruptions or noticeable variabilities. This state reflects the baseline performance where the operation proceeds as expected. The LV state includes scenarios where minor disruptions occurred, such as slight delays or minor deviations in precision. While these variabilities are noticeable, they do not significantly affect the overall operation or system performance. The MV state reflects situations where variabilities begin to impact the system more substantially. Examples include delays significant enough to disrupt schedules or minor incidents that require corrective actions but do not escalate into major issues. The HV state represents conditions where variabilities cause critical disruptions to the operation. This includes severe delays, major accidents, or incidents that significantly compromise the safety, efficiency, or overall integrity of the mooring process. These states are attributed proportionally to the observed data, reflecting their frequency and impact on the system.

BG_2 is an organizational function that encompasses a wide range of port management operations. Its role varies depending on the specific activities and requirements of the seaport, but it fundamentally oversees the overall management of the port by ensuring that operations are controlled, monitored, and efficiently coordinated. This function is critical for maintaining seamless day-to-day operations and adapting to the dynamic challenges of port environments. The variability in the performance of BG_2 is analysed by modelling its behaviour using a BN framework. This modelling approach, as described in Section 3.2.1, provides insights into how different organizational factors and conditions influence the effectiveness of BG_2, thereby supporting decision-making and performance optimization in seaport management.

BG_4 represents the logistics management and resource allocation capabilities of a seaport, encompassing its ability to supply and distribute necessary resources to various operational activities. This function is pivotal in ensuring that each section of the seaport operates efficiently. However, resource allocation is not uniform across all activities and depends on factors such as the level of investment, the priorities of stakeholders, and the criticality of specific operations to the overall performance of the seaport. In a technical context, resource prioritization is particularly important for high-stakes operations. For instance, logistics support for critical assets like quay cranes should be robust and well-structured. Quay cranes are integral to loading and unloading cargo, and any disruption in their operation can lead to significant delays, increased costs, and ripple effects throughout the supply chain. In contrast, yard-side operations, while essential, may not require the same level of resource intensity or redundancy because their disruptions, although impactful, are generally less immediate in their consequences. As a result, the output value of BG_4 is expected to vary across different functions, reflecting the unique logistical demands and resource priorities associated with each operation. However, for simplicity in this research, the UVI for all BG_4 elements is assigned the same value. After determining the performance variabilities of all background functions, which serve as inputs to the main model, and obtaining the SVI values for the foreground functions through the outlined approaches, the main model can now be computed. This allows for analysing and monitoring the impact of upstream variabilities on downstream functions. It is important to note that a comprehensive dataset, encompassing both objective and subjective information, was collected from a specific seaport. However, due to confidentiality agreements, the name of this seaport cannot be disclosed.

Figures 13, and 14 illustrate the performance variability values for all functions, including the SVI and UVI.

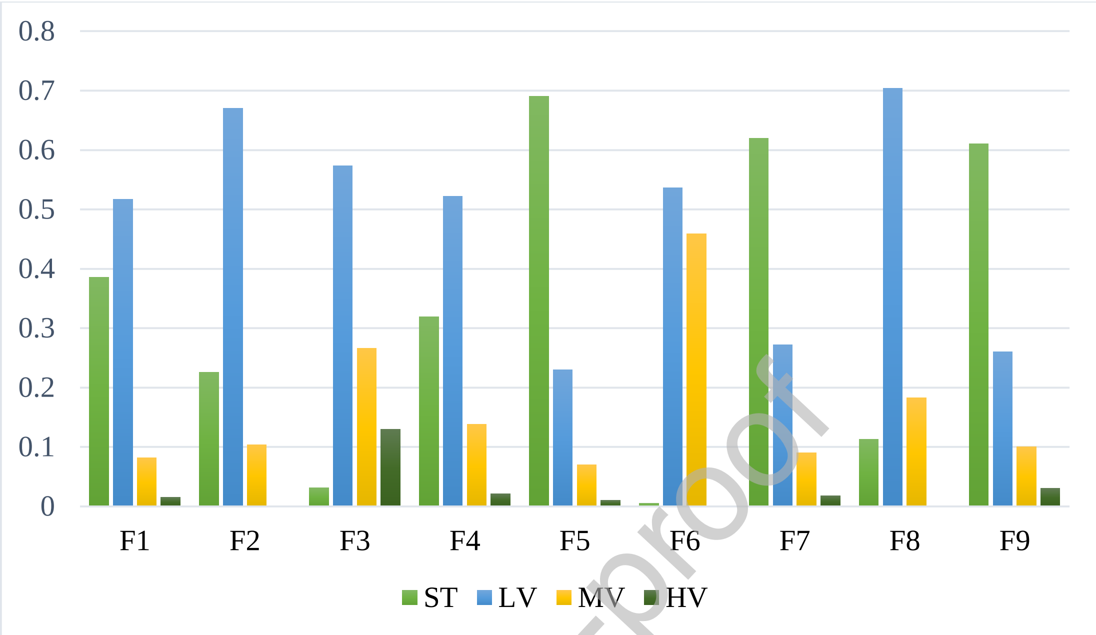

Figure 13: The SVI values.

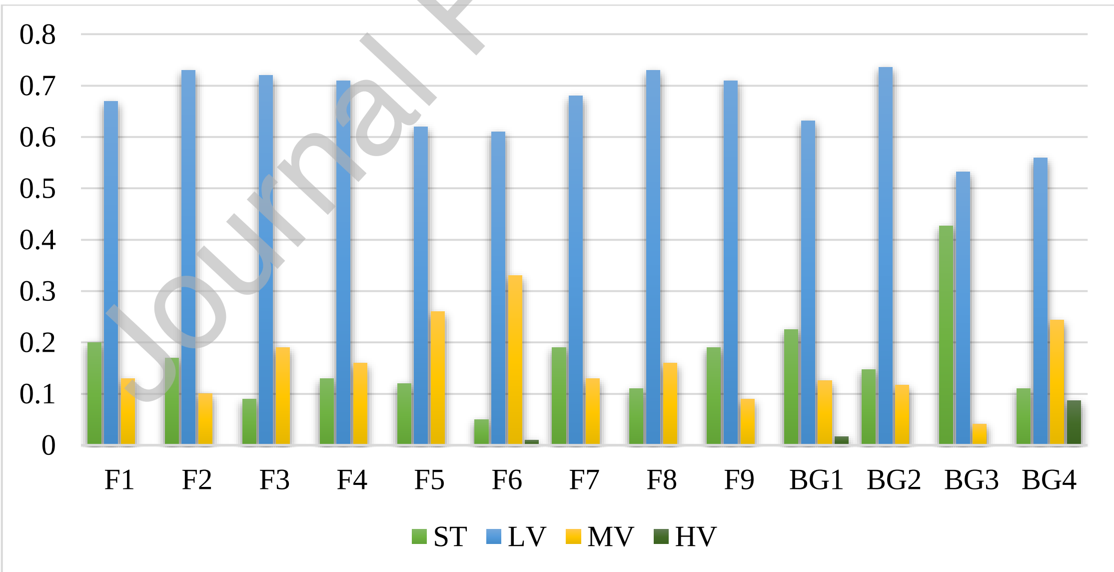

Figure 14: The UVI values.

4.4. Criticality matrix development

Once the performance variability for each function is quantified, the next step is to identify critical functions and evaluate the system's overall weaknesses from a systemic perspective. To achieve this, the UVI values are assigned appropriate scores, as outlined in Section 3.3.3, to derive a unique representative value for each function. This process involves calculating the mean, standard deviation, and the lower and upper bounds of the variability.

To represent the variability probabilistically, it is assumed that these aggregate scores follow a normal distribution. This assumption is common in probabilistic modelling, as the normal distribution effectively captures central tendencies (mean) and variability (standard deviation) *(Mitrani, 2008)*. Table 11 provides the representative output values for each function, reflecting the variability and its implications for the system. It is to be noted that the lower and upper bounds are determined at a 95% confidence level through MCSs, utilizing 100,000 iterations for precision.

Table 11: The functions representative output values for resonance analysis.

Function	Mean value	Standard deviation	Lower bound	Upper bound	Severity level
F1	1.930	0.570	0.812	3.048	Moderate
F2	1.930	0.515	0.921	2.939	Critical
F3	2.100	0.520	1.082	3.118	Moderate
F4	2.030	0.538	0.976	3.084	Moderate
F5	2.140	0.601	0.963	3.317	Critical
F6	2.300	0.574	1.174	3.426	Critical
F7	1.940	0.562	0.838	3.042	Moderate
F8	2.050	0.517	1.036	3.064	Moderate
F9	1.880	0.520	0.861	2.899	Moderate
BG1	1.935	0.644	0.673	3.197	Critical
BG2	1.970	0.513	0.965	2.975	Critical
BG3	1.614	0.565	0.507	2.721	Moderate
BG4	2.308	0.779	0.781	3.835	Critical

To assess the magnitude of variability in critical functions, their severity levels are also determined. However, accurately quantifying the magnitude of variability and its impact in terms of severity requires an independent study, as this step is crucial for understanding the consequential effects of disruptions in various elements of a CSTS. Given the scope of this study, we have relied on expert judgment to classify each function into three categories of severity: minor, moderate, and critical, as presented in Table 11.

Figure 15 illustrates the criticality matrix, which maps functions to their appropriate positions within the matrix. In this framework, the vertical axis reflects performance variability, with evenly distributed boundaries defined by the nature of each function in the seaport context. The proposed matrix offers flexibility for adaptation based on user-specific requirements, enabling its application to diverse systems of interest.

Figure 15: The criticality matrix for identifying critical functions in resonance analysis.

The criticality analysis reveals that all functions fall into level B, indicating minor levels of variability. While these variabilities are relatively low, they still have the potential to contribute to negative resonance, especially when interacting with moderately variable or interconnected functions. Such interactions can propagate risks throughout the system.

According to Safety-II principle, variability at level B can be viewed as an asset, as it arises from the adaptive adjustments necessary for everyday operations. However, the criticality matrix utilizes mean values derived from variability distributions to categorize variability into three levels. To incorporate uncertainty into risk-based decision-making, the upper and lower bounds of variability scores can provide a more nuanced understanding of the confidence in the mean score's placement within the matrix. For instance, if the upper bound is considered and indicates higher criticality, it could flag functions for further investigation even when the mean suggests a lower criticality level. Using this approach, functions such as F_2, F_5, F_6, F_7, BG_1, BG_2, and BG_4 would move to level C when upper bounds are applied. This shift indicates that these functions exceed acceptable thresholds and signal a need for immediate attention to mitigate the risk of negative resonance. As a practical example, if F_7 (Cargo transport to yard storage) were to fall within the red zone of the criticality matrix, this would signal the need for immediate intervention by terminal logistics teams. In such a situation, additional trucks would be deployed to avoid delays in container movement to the yard. Without timely action, performance variability in this function could propagate upstream to the quayside, increase variability in quay crane operations, and ultimately elevate the probability of terminal-wide disruption. By allocating redundant trucking capacity and addressing the issue at its source, variability can be contained, localized, and eliminated before it cascades into broader system instability. All in all, this approach rigorously prioritizes functions for safety countermeasures, emphasizing the need to reduce variability in their outputs. Addressing these criticalities pre-emptively can prevent negative resonance and ensure system stability, particularly in downstream processes.

4.5. Model validation process

As outlined in Section 3.4, multiple approaches are employed to validate the proposed model and its findings. For the HTA and FRAM, the validation process involved consultation with seven experts, each possessing at least 15 years of experience in seaport operations. These experts, with minor revisions, confirmed that the activities represented in the HTA and subsequently modelled in the FRAM, along with their structures and interconnections, accurately reflect the most significant and realistic activities observed in practice. Additionally, the results were partially benchmarked against other studies *(Cho et al., 2018; Darbra and Casal, 2004; John et al., 2014; Majumdar et al., 2022; Mitra et al., 2024; Yin et al., 2024)*. However, identifying and aligning with similar studies for comparison proved challenging due to the limited availability of directly comparable research and the complexity of matching findings.

In addition to the previously mentioned methods, sensitivity analysis was performed to validate the BN models. This process involved two sequential steps. First, the developed BNs for SVI evaluation were analysed as a partial validation of the overall model. Second, the FRAM-based BN models, which map the relationships between functions, were validated through sensitivity analysis. Using GeNIe software, a derivative-based sensitivity analysis was conducted, allowing the quantification of how changes in the BN's parameters influence the target nodes by calculating their derivatives. In this approach, the software uses mathematical and numerical techniques to compute the derivative of the posterior probability distribution of each target node with respect to each parameter. For instance, if *P(C/A)* represents the probability of a child node C given a parent node *A*, the derivative value is obtained as $\frac{\partial P(C)}{\partial P(A)}$, which quantifies how *P(C)* changes when *P(A)* is adjusted. Larger derivatives signify that even minor changes in a parameter have a substantial impact on the target node. By comparing derivatives across various parameters, the most influential ones can be identified.

As shown in Table 12, the three highest derivatives were selected along with their associated nodes as examples. It is important to note that these selections are based on the ST state of the target node. In other words, by setting the target node's state to ST, the most sensitive parameters were identified. Additionally, the variation range of the leaf node's ST state is reported, illustrating the span of possible values. For instance, within the technological function, setting environmental conditions to stable, MTTR to short, and ensuring an adequate level of redundancy is expected to contribute to the stability of performance variability. The interval values are centred around the original ST state values of the target node, fulfilling Principle 1 of sensitivity analysis.

To address Principle 2, the top three nodes, along with their relevant states, were subjected to a 10% increase in their values to observe the combined effect on the target node. For human functions, since the initial values for these three top nodes were at their maximum (100%), a 10% decrease was applied instead.

The results indicate that the posterior probabilities of the target node for technological, organizational, and F_2 functions shifted favourably toward the ST state, resulting in a corresponding reduction in performance variability as the ST values increased. In contrast, for human functions, the posterior probabilities leaned toward greater performance variability, with an increase in the MV values. This demonstrates that the collective impact of changes in the selected nodes on the target node's probabilities is consistently more significant than the impact of individual changes in each node, thereby validating Principle 2.

Table 12: The sensitivity analysis results.

Function	Node	State	Interval	Derivative	Prior prob.	Posterior prob.	Performance variability
Organizational	Authority gradient	Balanced	[0.258-0.419]	0.160	ST=0.38 LV=0.52 MV=0.08 HV=0.02	ST=0.43 LV=0.49 MV=0.07 HV=0.01	PV_1=1.74 PV_2=1.66 ΔP=-5%
	Span of control	Balanced	[0.258-0.419]	0.160			
	Communication effectiveness	Adequate	[0.255-0.410]	0.155			
Technological	Environmental condition	Stable	[0.483-0.820]	0.337	ST=0.69 LV=0.23 MV=0.07 HV=0.01	ST=0.75 LV=0.20 MV=0.05 HV=0.00	PV_1=1.39 PV_2=1.30 ΔP=-7%
	MTTR	Short	[0.582-0.729]	0.148			
	Redundancy	Adequate	[0.575-0.722]	0.146			
Human	QSO	$S_{9,4}$	[0.353-0.540]	0.192	ST=0.00 LV=0.54 MV=0.46 HV=0.00	ST=0.00 LV=0.45 MV=0.55 HV=0.00	PV_1=2.46 PV_2=2.55 ΔP=+4%
	APP	$S_{3,3}$	[0.378-0.540]	0.181			
	TAC	$S_{1,3}$	[0.404-0.540]	0.167			
F_2	Internal	ST	[0.165-0.173]	0.174	ST=0.17 LV=0.73 MV=0.10 HV=0.00	ST=0.22 LV=0.70 MV=0.08 HV=0.00	PV_1=1.93 PV_2=1.86 ΔP=-4%
	BG_1	ST	[0.164-0.172]	0.173			
	F_1	ST	[0.162-0.171]	0.172			

4.6. Implications

Based on the obtained results and the associated discussion, several implications can be drawn to support various seaport stakeholders, each benefiting from these insights from different operational and strategic perspectives, as outlined below:

1) Immediate control through the prioritisation of “level C” functions.

When the upper confidence bounds of the UVI distributions are considered, seven elements, including F2 (Initial Safety and Security Checks), F5 (Quay crane operation), F6 (Quay crane operator), F7 (Cargo Transport to Yard Storage), BG1 (Vessel securely moored), BG2 (Port operations management) and BG4 (Resource management) migrate from a minor variability “B” zone to the critical “C” zone of the matrix. This shows that apparently “well-behaved” functions can become risk amplifiers once uncertainty is acknowledged, so early safeguards must focus on these nodes before local variation resonates through the wider seaport system.

2) Operational stakeholders (pilots, berth masters, equipment maintainers).

F2 (cargo-handling coordination) and F5/F6 (quay-crane and yard operations) emerge as volatility hot-spots; actions such as dynamic berth planning and predictive maintenance will give the biggest risk-reduction pay-off.

3) Strategic management (port authority & terminal operators).

BG1 (mooring practice) and BG2 (organisational control) highlight that managerial routines and safety culture are as variable as front-line work; leadership should institutionalise continuous monitoring and learning loops.

4) Logistics partners & investors.

BG4’s high variability underlines that resource-allocation policy (e.g., spare-part inventories, redundancy levels) directly drives systemic stability.

5) Resource-allocation rules derived from sensitivity analysis.

Derivative-based sensitivity reveals that keeping environmental conditions stable, MTTR short and redundancy adequate produces 5-7% shifts of the posterior toward the ST state for technological functions, whereas poor attention to these parameters moves human functions toward the MV state. This quantifies how marginal investment in redundancy or faster repair capability suppresses resonance potential system wide.

6) Balancing Safety-II adaptability with resonance prevention.

Although most mean UVIs sit in the “beneficial variability” band (level B), the wide upper tails caution against complacency; variability is an asset only while resources exist to damp it. The integrated FRAM-BN model makes that trade-off explicit by letting analysts toggle between mean, lower- and upper-bound scenarios during what-if simulations.

7) Methodological generalisation to other complex transport systems.

The quantitative FRAM-BN coupling used here aligns with the recent trend in maritime-risk science toward data-driven Bayesian networks combined with functional models *(Guo et al., 2023; Mohsendokht et al., 2024a, 2024b; Z. Yang et al., 2025).*

It is noted that the proposed framework has been designed to address the inherent complexity of safety assessment in complex socio-technical systems through a Safety-II-oriented perspective. While seaport operations were selected as the primary case study to demonstrate applicability and practical feasibility, the methodological structure is not domain specific. Owing to its modular design principles and emphasis on functional variability, the framework is readily generalisable to other complex sectors, such as aviation, rail, energy, and healthcare, where multi-actor interactions, dynamic operating environments, and emergent performance conditions similarly influence system resilience and safety outcomes. With appropriate contextualisation of system functions and domain-specific variables, these sectors can likewise adopt the framework to undertake systemic, performance-based Safety-II risk assessments. The applicability of this approach can be considered from two perspectives, reflecting both the commonalities and the sector-specific differences.

From a commonality perspective, these sectors share fundamental characteristics with maritime operations, including high interdependence among human, technical, and organizational elements, reliance on continuous coordination between multiple agents, and the presence of tightly coupled subsystems where small performance variabilities can lead to potential negative resonances with significant system-wide consequences. In aviation and rail transport, as in seaports, operational safety depends on synchronized human-machine interactions, adherence to procedural constraints, and resilience to unexpected disturbances. Similarly, healthcare systems exhibit comparable socio-technical complexity, where variability in human performance and resource constraints can critically affect outcomes. Thus, the framework's focus on modelling functional variability and emergent behaviour makes it well suited for analysing these domains.

From a uniqueness perspective, each sector exhibits distinct features that require contextual adaptation of the framework. For instance, aviation and rail industries often possess richer operational and safety data through advanced monitoring systems and regulatory reporting, which can reduce subjective bias in quantification of performance variability and improve the empirical grounding of probabilistic models. In contrast, healthcare environments are characterized by higher contextual diversity and limited standardization, meaning that qualitative judgment and expert elicitation remain essential for capturing functional dependencies and performance variability. Consequently, appropriate contextualization of system functions, data availability, and performance indicators will be essential when adapting the framework to each domain.

In summary, the proposed Safety-II-based framework provides a flexible foundation for systemic and performance-based risk assessment across diverse complex socio-technical sectors. Its modular architecture allows for both cross-domain generalization and domain-specific customization, ensuring its applicability to different CSTS.

5. CONCLUSION

In this study, a novel systemic risk assessment approach is designed to capture the dynamic interactions among the various elements of a seaport. Performance variability is acknowledged as a distinctive framework for expressing and understanding the interdependencies between diverse functions. The FRAM serves as the foundational component of the approach, enabling the visualization of real-world relationships between activities, referred to as functions, within a seaport context. To enhance FRAM's capability for quantitative analysis, it is integrated with BN, allowing consideration of both internal and external factors that may influence individual functions. The proposed methodology builds upon the principles of the Safety-II concept, emphasizing a functional safety perspective. The outcomes of the study and the application of the framework provide deeper insights into system dynamics and offer more practical, versatile strategies for improving overall system safety.

Given the obtained results, insights, and implications, this study makes several significant contributions, as follows:

1) Holistic analysis of CSTS elements: Technological, human, and organizational functions within the CSTS are analysed to evaluate their internal and external performance variabilities, considering factors such as operational uncertainties, environmental conditions, and human performance fluctuations.
2) Function interactions: Interactions between functions are systematically analysed to track upstream-downstream performance variability, assessing their impacts on the overall system. This approach highlights critical dependencies and emergent behaviours.
3) Comprehensive risk analysis: The framework supports both retrospective and prospective evaluations of performance variability, providing actionable insights for addressing risks effectively.
4) Enhanced decision-making: By quantifying and visualizing performance variabilities, the framework enables risk-based decision-making, helping prioritize interventions and implement targeted risk management measures.

This integrated approach provides a solid foundation for understanding and mitigating systemic risks in complex socio-technical system environments. Nevertheless, while the framework demonstrates strong potential for comprehensive systemic risk assessment, several limitations and avenues for future enhancement remain.

First, the FRAM model development was based on expert knowledge and focused on key operational functions. As system size and complexity grow, the number of functions and their interdependencies may expand significantly, making manual modelling increasingly demanding and time-consuming. Future research could integrate machine-learning-assisted techniques, such as those informed by HTA analysis, to support automated function identification and coupling detection. These techniques would serve as an advisory tool to assist experts, thereby improving scalability and modelling efficiency while preserving domain oversight. Second, limited availability of empirical data for human and organisational functions necessitated reliance on expert judgment. Although expert elicitation remains a widely accepted practice in complex socio-technical analyses where datasets are scarce, this reliance may introduce subjectivity and uncertainty. In particular, access to verifiable performance-related information for organisational functions is often constrained, as managers and policymakers may be reluctant to critically examine or disclose internal performance practices. Consequently, unlike technological functions, where structured monitoring and measurable performance data are routinely available, organisational and human performance data remain largely qualitative and under-reported. Furthermore, while certain industries such as nuclear energy sector have established quantitative data collection frameworks for human reliability, similar systematic mechanisms are still emerging within the maritime sector, especially for capturing performance variability rather than traditional error-based measures. To address these gaps, future works should focus on developing structured platforms for continuous collection and documentation of operational performance data, covering both routine and non-routine conditions. Such efforts would reduce reliance on subjective input, enhance traceability, and enable more robust, data-driven modelling of human and organisational variability.

APPENDICES

Appendix A

Regarding the application of modified CREAM methodology to assess the SVI of human functions in step 3 the following example is brought up here to clarify the procedure. Let's say, in rule number k, out of the nine CPCs, four have positive effects, three have negative effects, and two have neutral effects. Referring to the vertical axis of the diagram, which corresponds to the value four, and examining the shares of the slots associated with different, it is evident that there is one block for "strategic", five blocks for "tactical", and none for the other CCMs. Based on this, β^+ is estimated as:

$$\beta^+ =\{(0.17,CCM_1), (0.73,CCM_2), (0, CCM_3), (0,CCM_4)\}$$

Similarly, using the horizontal axis of the diagram and identifying the value three, two blocks are observed for "opportunistic", five blocks for "tactical", and none for the other CCMs. Consequently, β^- is calculated as:

$$\beta^- =\{(0,CCM_1), (0.71,CCM_2), (0.29, CCM_3), (0,CCM_4)\}$$

It should be emphasized that the "neutral" effect does not contribute to the integrated result, as it has already been accounted for in the uncertainty, and its belief degree is therefore excluded from the process.

Once the positive and negative belief degrees are determined, and the weights derived using the AHP approach are incorporated, evidential reasoning is employed to synthesize this information. This process delivers the final combined belief degree for each CCM.

Appendix B

Regarding the DSET approach, let's consider a set of *n* mutually exclusive and exhaustive propositions, referred to in this context as the BN states, $\Omega=\{X_0, X_1,\dots, X_n\}$. This set Ω is called the frame of discernment.

The power set, denoted 2^Ω, includes all possible subsets of Ω, including the empty set $(\emptyset)$ and Ω itself. For a frame $\Omega=\{X_0, X_1\}$, the power set is: $2^\Omega=\{\emptyset,\{X_0\},\{X_1\},\{X_0,X_1\}\}$. In general, for n elements in Ω, 2^n subsets are formed.

DSET comprises three vital functions: the Basic Probability Assignment (BPA), the Belief Function (BEL), and Plausibility Function (PL). BPA, denoted as $m(A)$, assigns a mass of probability to a subset A of the frame of discernment Ω, where $A\in\Omega$.

The following rules are applied: The mass of the empty set is always zero: m(∅)=0, and the sum of all masses over 2^{Ω} is 1, which is illustrated as $\sum_{A\in\Omega} m(A) = 1$. *A* is referred to as a focal element if *m(A)>0* and *m(A)* represents the extent to which the evidence supports the proposition *A*.

Continuing, the BEL serves as the lower bound of the probability interval, while the PL acts as the upper bound. They are defined as follows:

$$BEL\ (X) = \sum_{P\subseteq X} \prod_{1\leq i\leq n} m_i(P_i) \quad (1)$$

$$PL(X) = 1 - BEL(\bar{X}) \quad (2)$$

where *P* is the proper subset of the set of interest *(X)*, *i* is the number of beliefs, and $\bar{X}$ signifies the complement of *X*, indicating that the belief is governed by the principle that the total basic probability BPA must equal 1.

When multiple pieces of evidence from different sources are presented, the fusion of beliefs is determined by the combination rule of DSET as follows:

$$m(A) = m_1(A) \oplus m_2(A) \ldots m_n(A) = \frac{\sum_{B\cap C\cap \ldots Z=A} m_1(B) m_2(C) \ldots m_n(Z)}{1-K} \quad (3)$$

when *A≠∅*, *m(∅)=0*,

and where,

$$K = \sum_{B\cap C\cap \ldots Z=\emptyset} m_1(B) m_2(C) \ldots m_n(Z) \quad (4)$$

K represents the level of conflict between the pieces of evidence, with *K=0* indicating no conflict and *K=1* signifying complete contradiction between the evidence.

Appendix C

Table I: Symbols and definitions used in the proposed methodology.

Symbol	*Description*
ST	Stable conditions, no variability
LV	Low variability
MV	Moderate variability
HV	High variability
CPT	Conditional Probability Table
CPC	Common Performance Condition
P(C/R)	Conditional probability of child node state C given parent node state *R*
P(CCM)	Probability of performance variability in terms of Contextual Control Mode
$S_{i,j}$	The state of CPC i and number j
$P(S_{i,j})$	Probability of specific states of a given common performance condition
$P(CCM/S_{i,j})$	Conditional probability of performance variability given a CPC states
$\frac{\partial P(C)}{\partial P(A)}$	The derivative value of child node C given a parent node A
MTTF	Mean Time To Failure
MTTR	Mean Time To Repair
SVI	Self-contained Variability Index
UVI	Upstream Variability Index
β^+	Belief degrees with positive effect
β^-	Belief degrees with negative effect
β	Combinatory belief degrees
W^+	The corresponding weights of CPCs with positive effects
W^-	The corresponding weights of CPCs with negative effects

F_i	Foreground function number i
BG_i	Background function number i

Table II: The expert's profile and their related experience and expertise.

Number	Title	Educational level	Experience (years)	Location	Specialization
1	General Manager	MSc	20	Iran	Port master planning; concession/PPP contract management; stakeholder engagement; ESG & sustainability; business continuity & resilience.
2	Operations Manager	MSc	15	Iran	Berth planning & vessel scheduling; quay-crane assignment; yard planning & resource optimization; stowage coordination with shipping lines.
3	Operations Manager	Ph.D	12	Canada	Container terminal management; stevedoring planning; turnaround-time optimization.
4	HSE Director	Ph.D	10	Belgium	ISO 45001/14001 systems; HAZID/HAZOP/JSA risk assessment; emergency response & oil-spill (ICS) planning; contractor HSE auditing.
5	Harbour Master	MSc	18	USA	VTS & navigational safety; pilotage & towage coordination; mooring/lines safety; ISPS drills & security interface; incident investigation & root-cause analysis.
6	Port Planning	MSc	16	UK	Berth/yard capacity modelling; approach-channel design & navigational risk; asset management (PIANC/ICE standards).
7	Terminal Systems & Automation Manager	MSc	14	Australia	TOS configuration; yard optimization & equipment dispatching (ASC/RTG/AGV); EDI/port community systems; operational analytics & dashboarding.

Table III: CPC estimates based on expert elicitation.

CPC	States	Expert 1	Expert 2	Expert 3	Aggregated value	Effects on performance variability
TAC	$S_{1,1}$	0.00	0.00	0.00	0.0000	Negative
	$S_{1,2}$	0.10	0.15	0.10	0.0022	Neutral
	$S_{1,3}$	0.90	0.85	0.90	0.9978	Positive
HMI	$S_{2,1}$	0.00	0.00	0.00	0.0000	Negative
	$S_{2,2}$	0.10	0.00	0.05	0.0000	Neutral
	$S_{2,3}$	0.75	0.80	0.85	0.9941	Neutral
	$S_{2,4}$	0.15	0.20	0.10	0.0059	Positive
APP	$S_{3,1}$	0.00	0.00	0.00	0.0000	Negative
	$S_{3,2}$	0.10	0.00	0.05	0.0000	Neutral

	$S_{3,3}$	0.90	1.00	0.95	1.0000	Positive
COW	$S_{4,1}$	0.35	0.25	0.30	0.0714	Negative
	$S_{4,2}$	0.65	0.75	0.70	0.9286	Neutral
	$S_{4,3}$	0.00	0.00	0.00	0.0000	Positive
NGC	$S_{5,1}$	0.25	0.10	0.20	0.0092	Negative
	$S_{5,2}$	0.75	0.90	0.80	0.9908	Neutral
	$S_{5,3}$	0.00	0.00	0.00	0.0000	Positive
ATT	$S_{6,1}$	0.30	0.30	0.20	0.0542	Negative
	$S_{6,2}$	0.70	0.60	0.75	0.9458	Neutral
	$S_{6,3}$	0.00	0.10	0.05	0.0000	Positive
CRS	$S_{7,1}$	0.05	0.05	0.10	0.0007	Negative
	$S_{7,2}$	0.70	0.75	0.75	0.9806	Neutral
	$S_{7,3}$	0.25	0.20	0.15	0.0187	Positive
TCQ	$S_{8,1}$	0.00	0.00	0.00	0.0000	Negative
	$S_{8,2}$	0.15	0.10	0.10	0.0030	Neutral
	$S_{8,3}$	0.80	0.75	0.85	0.9963	Neutral
	$S_{8,4}$	0.05	0.15	0.05	0.0007	Positive
QSO	$S_{9,1}$	0.05	0.00	0.05	0.0000	Negative
	$S_{9,2}$	0.25	0.30	0.35	0.0819	Negative
	$S_{9,3}$	0.70	0.70	0.60	0.9181	Neutral
	$S_{9,4}$	0.00	0.00	0.00	0.0000	Positive

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENTS

This work is supported by the European Research Council project (TRUST CoG 2019 864724).

REFERENCES

Adhita, I.G.M.S., Fuchi, M., Konishi, T., Fujimoto, S., 2023. Ship navigation from a Safety-II perspective: A case study of training-ship operation in coastal area. Reliab. Eng. Syst. Saf. 234, 109140. https://doi.org/10.1016/j.ress.2023.109140

Aeon, B., Faber, A., Panaccio, A., 2021. Does time management work? A meta-analysis. PLOS ONE 16, e0245066. https://doi.org/10.1371/journal.pone.0245066

Ahmed, F., Robinson, S., Tako, A.A., 2014. Using the structred analysis and design technique (SADT) in simulation conceptual modeling, in: Proceedings of the Winter Simulation Conference 2014. Presented at the 2014 Winter Simulation Conference - (WSC 2014), IEEE, Savanah, GA, USA, pp. 1038–1049. https://doi.org/10.1109/WSC.2014.7019963

Aven, T., 2022. A risk science perspective on the discussion concerning Safety I, Safety II and Safety III. Reliab. Eng. Syst. Saf. 217, 108077. https://doi.org/10.1016/j.ress.2021.108077

Baxter, G., Sommerville, I., 2011. Socio-technical systems: From design methods to systems engineering. Interact. Comput. 23, 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

Bayramova, A., Edwards, D.J., Roberts, C., Rillie, I., 2023. Enhanced safety in complex socio-technical systems via safety-in-cohesion. Saf. Sci. 164, 106176. https://doi.org/10.1016/j.ssci.2023.106176

Birolini, A., 2017. Reliability Engineering: Theory and Practice. Springer.

Busch, T., 2011. Organizational adaptation to disruptions in the natural environment: The case of climate change. Scand. J. Manag. 27, 389–404. https://doi.org/10.1016/j.scaman.2010.12.010

Bytheway, C.W., 2007. FAST creativity and innovation: Rapidly Improving Processes, Product Development and Solving Complex Problems. J. Ross Publishing.
Cantelmi, R., Steen, R., Gravio, G.D., Patriarca, R., 2025. An explorative Bayesian analysis of functional dependencies in emergency management systems. Syst. Eng. 28, 82–99. https://doi.org/10.1002/sys.21783
Carlo, H.J., Vis, I.F.A., Roodbergen, K.J., 2015. Seaside operations in container terminals: literature overview, trends, and research directions. Flex. Serv. Manuf. J. 27, 224–262. https://doi.org/10.1007/s10696-013-9178-3
Carmeli, A., Tishler, A., 2004. Resources, capabilities, and the performance of industrial firms: A multivariate analysis. Manag. Decis. Econ. 25, 299–315. https://doi.org/10.1002/mde.1192
Catelani, M., Ciani, L., Guidi, G., Patrizi, G., 2021. An enhanced SHERPA (E-SHERPA) method for human reliability analysis in railway engineering. Reliab. Eng. Syst. Saf. 215, 107866. https://doi.org/10.1016/j.ress.2021.107866
Ceylan, B.O., Akyuz, E., Arslan, O., 2021. Systems-Theoretic Accident Model and Processes (STAMP) approach to analyse socio-technical systems of ship allision in narrow waters. Ocean Eng. 239, 109804. https://doi.org/10.1016/j.oceaneng.2021.109804
Chinenye Gbemisola Okatta, Funmilayo Aribidesi Ajayi, Olufunke Olawale, 2024. ENHANCING ORGANIZATIONAL PERFORMANCE THROUGH DIVERSITY AND INCLUSION INITIATIVES: A META-ANALYSIS. Int. J. Appl. Res. Soc. Sci. 6, 734–758. https://doi.org/10.51594/ijarss.v6i4.1065
Cho, H.S., Lee, J.S., Moon, H.C., 2018. Maritime Risk in Seaport Operation: A Cross-Country Empirical Analysis with Theoretical Foundations. Asian J. Shipp. Logist. 34, 240–247. https://doi.org/10.1016/j.ajsl.2018.09.010
Clarotti, C., Lannoy, A., Odin, S., Procaccia, H., 2004. Detection of equipment aging and determination of the efficiency of a corrective measure. Reliab. Eng. Syst. Saf. 84, 57–64. https://doi.org/10.1016/j.ress.2004.01.005
Cooper, M.D., 2022. The Emperor has no clothes: A critique of Safety-II. Saf. Sci. 152, 105047. https://doi.org/10.1016/j.ssci.2020.105047
Costa, L.Q.M.D., Cavalcante, C.A.V., 2022. A REVIEW ON THE STUDY OF MAINTENANCE EFFECTIVENESS. Pesqui. Oper. 42, e263613. https://doi.org/10.1590/0101-7438.2022.042nspe1.00263613
Darbra, R.-M., Casal, J., 2004. Historical analysis of accidents in seaports. Saf. Sci. 42, 85–98. https://doi.org/10.1016/S0925-7535(03)00002-X
Deng, Z.Z.W., 2000. Crane Design Manual(Chinese Edition). ZHANG ZHI WEN DENG: 9787113025717 - AbeBooks. https://www.abebooks.co.uk/9787113025717/Crane-Design-ManualChinese-Edition-ZHANG-7113025714/plp.
Diem Le, C.T., Pakurár, M., Kun, I.A., Oláh, J., 2023. Correction: The impact of factors on information sharing: An application of meta-analysis. PLOS ONE 18, e0296555. https://doi.org/10.1371/journal.pone.0296555
Díez, F.J., 1993. Parameter adjustment in Bayes networks. The generalized noisy OR–gate, in: Uncertainty in Artificial Intelligence. Elsevier, pp. 99–105. https://doi.org/10.1016/B978-1-4832-1451-1.50016-0
Diez, F.J., Druzdzel, M.J., 2006. Canonical probabilistic models for knowledge engineering. Technical report, UNED. http://www.cisiad.uned.es.
Donelson, D.C., Garfinkel, J.A., Roudini, S., 2023. Regulatory Policy Enforcement and Corporate Performance. SSRN Electron. J. https://doi.org/10.2139/ssrn.4388696
Duan, C., Deng, T., Song, L., Wang, M., Sheng, B., 2023. An adaptive reliability-based maintenance policy for mechanical systems under variable environments. Reliab. Eng. Syst. Saf. 238, 109396. https://doi.org/10.1016/j.ress.2023.109396
Erik, H., 2017. FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-technical Systems, 1st ed. CRC Press. https://doi.org/10.1201/9781315255071
Ferreira, R.J.P., De Almeida, A.T., Cavalcante, C.A.V., 2009. A multi-criteria decision model to determine inspection intervals of condition monitoring based on delay time analysis. Reliab. Eng. Syst. Saf. 94, 905–912. https://doi.org/10.1016/j.ress.2008.10.001

Germain, R., Claycomb, C., Dröge, C., 2008. Supply chain variability, organizational structure, and performance: The moderating effect of demand unpredictability. J. Oper. Manag. 26, 557–570. https://doi.org/10.1016/j.jom.2007.10.002

Grabbe, N., Arifagic, A., Bengler, K., 2022. Assessing the reliability and validity of an FRAM model: the case of driving in an overtaking scenario. Cogn. Technol. Work 24, 483–508. https://doi.org/10.1007/s10111-022-00701-7

Gros, X.E., 1997. Data Fusion – A Review, in: NDT Data Fusion. Elsevier, pp. 5–42. https://doi.org/10.1016/B978-034067648-6/50004-9

Grossman, R., Nolan, K., Rosch, Z., Mazer, D., Salas, E., 2022. The team cohesion-performance relationship: A meta-analysis exploring measurement approaches and the changing team landscape. Organ. Psychol. Rev. 12, 181–238. https://doi.org/10.1177/20413866211041157

Guo, Y., Jin, Y., Hu, S., Yang, Z., Xi, Y., Han, B., 2023. Risk evolution analysis of ship pilotage operation by an integrated model of FRAM and DBN. Reliab. Eng. Syst. Saf. 229, 108850. https://doi.org/10.1016/j.ress.2022.108850

Haas, J., 2016. Gard guidance on freight containers. Gard, ISBN: 978-82-90344-35-6.

Ham, D.-H., Park, J., 2020. Use of a big data analysis technique for extracting HRA data from event investigation reports based on the Safety-II concept. Reliab. Eng. Syst. Saf. 194, 106232. https://doi.org/10.1016/j.ress.2018.07.033

Hasan, S., Ali, M., Kurnia, S., Thurasamy, R., 2021. Evaluating the cyber security readiness of organizations and its influence on performance. J. Inf. Secur. Appl. 58, 102726. https://doi.org/10.1016/j.jisa.2020.102726

Henrion, M., 1989. Some Practical Issues in Constructing Belief Networks. in: Uncertainty in Artificial Intelligence. Elsevier Science Publisher B.V, North Holland, pp. 161-174.

Hillman, A.J., Keim, G.D., 2001. Shareholder Value, Stakeholder Management, and Social Issues: What's the Bottom Line? Strategic Management Journal, 22(2), 125–139. http://www.jstor.org/stable/3094310.

Hirose, T., Sawaragi, T., 2020. Extended FRAM model based on cellular automaton to clarify complexity of socio-technical systems and improve their safety. Saf. Sci. 123, 104556. https://doi.org/10.1016/j.ssci.2019.104556

Hirose, T., Sawaragi, T., 2019. Development of FRAM Model Based on Structure of Complex Adaptive Systems to Visualize Safety of Socio-Technical Systems. IFAC-Pap. 52, 13–18. https://doi.org/10.1016/j.ifacol.2019.12.075

Hollnagel, E., 2018. Safety–I and Safety–II: The Past and Future of Safety Management, 1st ed. CRC Press. https://doi.org/10.1201/9781315607511

Hollnagel, E., 1998. Cognitive Reliability and Error Analysis Method (CREAM). Elsevier. https://doi.org/10.1016/B978-0-08-042848-2.X5000-3

Hollnagel, E., Rees Hill, Slater, D., 2023. The FRAM model Visualiser (FMV)-The Web Version. https://doi.org/10.13140/RG.2.2.23719.32162

Hollnagel, E., Wears, R.L., Braithwaite, J., 2015. From Safety-I to Safety-II: A White Paper. https://doi.org/10.13140/RG.2.1.4051.5282

Huang, Q., Wei, Y., Zhang, J., Xu, X., Jin, X., 2025. A comprehensive CREAM method for human reliability analysis of armored vehicle crews based on cognitive performance and operational environment. Risk Anal. risa.70050. https://doi.org/10.1111/risa.70050

Jensen, A., Aven, T., 2018. A new definition of complexity in a risk analysis setting. Reliab. Eng. Syst. Saf. 171, 169–173. https://doi.org/10.1016/j.ress.2017.11.018

Jo, J.-H., Kim, S., 2019. Key Performance Indicator Development for Ship-to-Shore Crane Performance Assessment in Container Terminal Operations. J. Mar. Sci. Eng. 8, 6. https://doi.org/10.3390/jmse8010006

John, A., Paraskevadakis, D., Bury, A., Yang, Z., Riahi, R., Wang, J., 2014. An integrated fuzzy risk assessment for seaport operations. Saf. Sci. 68, 180–194. https://doi.org/10.1016/j.ssci.2014.04.001

Jones, B., Jenkinson, I., Yang, Z., Wang, J., 2010. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliab. Eng. Syst. Saf. 95, 267–277. https://doi.org/10.1016/j.ress.2009.10.007

Kannally, C., Paladugu, A., Nijveldt, R., McSherry, L., IJtsma, M., 2025. An Exploratory Study of Contextual Control Modes in Teamwork. Hum. Factors J. Hum. Factors Ergon. Soc. 67, 409–426. https://doi.org/10.1177/00187208241292669

Kaya, G.K., Ozturk, F., Sariguzel, E.E., 2021. System-based risk analysis in a tram operating system: Integrating Monte Carlo simulation with the functional resonance analysis method. Reliab. Eng. Syst. Saf. 215, 107835. https://doi.org/10.1016/j.ress.2021.107835

Kim, Y.C., Yoon, W.C., 2021. Quantitative representation of the functional resonance analysis method for risk assessment. Reliab. Eng. Syst. Saf. 214, 107745. https://doi.org/10.1016/j.ress.2021.107745

Kołowrocki, K., Soszyńska-Budny, J., 2011. Reliability and safety of complex technical systems and processes: Modeling – Identification – Prediction - Optimization. Springer.

Konstandinidou, M., Nivolianitou, Z., Kiranoudis, C., Markatos, N., 2006. A fuzzy modeling application of CREAM methodology for human reliability analysis. Reliab. Eng. Syst. Saf. 91, 706–716. https://doi.org/10.1016/j.ress.2005.06.002

Kuai, Y., Wang, H., 2025. Geopolitical Risk and Corporate ESG Performance: Evidence from China. Emerg. Mark. Finance Trade 61, 1010–1029. https://doi.org/10.1080/1540496X.2024.2404176

Lee, J., Chung, H., 2018. A new methodology for accident analysis with human and system interaction based on FRAM: Case studies in maritime domain. Saf. Sci. 109, 57–66. https://doi.org/10.1016/j.ssci.2018.05.011

Leveson, N., 2004. A new accident model for engineering safer systems. Saf. Sci. 42, 237–270. https://doi.org/10.1016/S0925-7535(03)00047-X

Li, P., Chen, G., Dai, L., Zhang, L., 2012. A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks. Saf. Sci. 50, 1569–1583. https://doi.org/10.1016/j.ssci.2012.03.017

Li, W., Zhang, L., Liang, W., 2017. An Accident Causation Analysis and Taxonomy (ACAT) model of complex industrial system from both system safety and control theory perspectives. Saf. Sci. 92, 94–103. https://doi.org/10.1016/j.ssci.2016.10.001

Liu, L., 2021. The Politics of (No) Compromise: Information Acquisition, Policy Discretion, and Reputation. https://doi.org/10.48550/ARXIV.2111.00522

Lu, B., Zhou, X., 2019. Quality and reliability oriented maintenance for multistage manufacturing systems subject to condition monitoring. J. Manuf. Syst. 52, 76–85. https://doi.org/10.1016/j.jmsy.2019.04.003

LÜScher, L.S., Lewis, M.W., 2008. Organizational Change and Managerial Sensemaking: Working Through Paradox. Acad. Manage. J. 51, 221–240. https://doi.org/10.5465/amj.2008.31767217

Luva, B., Naweed, A., 2024. King of the castle: organisational influences on authority gradients between network controllers and other team members. Ergonomics 67, 34–49. https://doi.org/10.1080/00140139.2023.2202844

Majumdar, A., Manole, I., Nalty, R., 2022. Analysis of Port Accidents and Calibration of Heinrich's Pyramid. Transp. Res. Rec. J. Transp. Res. Board 2676, 476–489. https://doi.org/10.1177/03611981211044447

Martinetti, A., Chatzimichailidou, M.M., Maida, L., Van Dongen, L., 2019. Safety I–II, resilience and antifragility engineering: a debate explained through an accident occurring on a mobile elevating work platform. Int. J. Occup. Saf. Ergon. 25, 66–75. https://doi.org/10.1080/10803548.2018.1444724

Martins, J.B., Carim, G., Saurin, T.A., Costella, M.F., 2022. Integrating Safety-I and Safety-II: Learning from failure and success in construction sites. Saf. Sci. 148, 105672. https://doi.org/10.1016/j.ssci.2022.105672

Mees, B., McMurray, A., Chhetri, P., 2016. Organisational resilience and emergency management. Australian Journal of Emergency Management, 31(2), 38–43. http://ecite.utas.edu.au/128982/1/128982%20-%20Organisational%20resilience%20and%20emergency%20management.pdf.

Melchers, R.E., 2005. The effect of corrosion on the structural reliability of steel offshore structures. Corros. Sci. 47, 2391–2410. https://doi.org/10.1016/j.corsci.2005.04.004

Mitra, A., Youdon, C., Chauhan, P., Shaw, R., 2024. Systemic risk capability assessment methodology: A new approach for evaluating inter-connected risks in seaport ecosystems. Prog. Disaster Sci. 22, 100325. https://doi.org/10.1016/j.pdisas.2024.100325
Mitrani, I., 2008. Probabilistic Modelling. Cambridge University Press; 2nd edition. 978-0521585309.
Mohaghegh, Z., Kazemi, R., Mosleh, A., 2009. Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. Reliab. Eng. Syst. Saf. 94, 1000–1018. https://doi.org/10.1016/j.ress.2008.11.006
Mohsendokht, M., Jamshidi, M., 2021. Re-evaluation of level 2 PSA results of a VVER-1000 NPP with consideration of innovative severe accident management strategies. Ann. Nucl. Energy 151, 107966. https://doi.org/10.1016/j.anucene.2020.107966
Mohsendokht, M., Kontovas, C., Chang, C.-H., Qu, Z., Li, H., Yang, Z., 2025. Resilience analysis of seaports: a critical review of development and research directions. Marit. Policy Manag. 1–36. https://doi.org/10.1080/03088839.2025.2483410
Mohsendokht, M., Li, H., Kontovas, C., Chang, C., Qu, Z., Yang, Z., 2024a. Enhancing maritime transportation security: A data-driven Bayesian network analysis of terrorist attack risks. Risk Anal. risa.15750. https://doi.org/10.1111/risa.15750
Mohsendokht, M., Li, H., Kontovas, C., Chang, C.-H., Qu, Z., Yang, Z., 2024b. Decoding dependencies among the risk factors influencing maritime cybersecurity: Lessons learned from historical incidents in the past two decades. Ocean Eng. 312, 119078. https://doi.org/10.1016/j.oceaneng.2024.119078
Moulahi, M.H., Zdiri, S., 2025. Monitoring Degradation Process Level and Reliability Improvement based on Decision Making with GLR Test. J. Syst. Sci. Syst. Eng. https://doi.org/10.1007/s11518-025-5658-8
Musheke, M.M., Phiri, J., 2021. The Effects of Effective Communication on Organizational Performance Based on the Systems Theory. Open J. Bus. Manag. 09, 659–671. https://doi.org/10.4236/ojbm.2021.92034
Nasim, M.A., Yadav, R.S., Dash, S.S., Bamel, U., 2023. Leadership styles and safety culture – a meta-analytic study. Int. J. Organ. Anal. 31, 3233–3250. https://doi.org/10.1108/IJOA-02-2022-3166
Natsir, I., Ode, H., Irfana, T.B., Pulungan, D.R., Sihite, M., 2024. An Analysis of the Role of Human Resource Management in Enhancing Organizational Efficiency and Effectiveness. Int. J. Econ. IJEC 3, 1089–1097. https://doi.org/10.55299/ijec.v3i2.1097
Noor Arzahan, I.S., Ismail, Z., Yasin, S.M., 2022. Safety culture, safety climate, and safety performance in healthcare facilities: A systematic review. Saf. Sci. 147, 105624. https://doi.org/10.1016/j.ssci.2021.105624
Ozdemir, S., Carlos Fernandez De Arroyabe, J., Sena, V., Gupta, S., 2023. Stakeholder diversity and collaborative innovation: Integrating the resource-based view with stakeholder theory. J. Bus. Res. 164, 113955. https://doi.org/10.1016/j.jbusres.2023.113955
Papadimitriou, E., Pooyan Afghari, A., Tselentis, D., Van Gelder, P., 2022. Road-safety-II: Opportunities and barriers for an enhanced road safety vision. Accid. Anal. Prev. 174, 106723. https://doi.org/10.1016/j.aap.2022.106723
Park, J., Kim, Ji-tae, Lee, S., Kim, Jonghyun, 2018. Modeling Safety-II based on unexpected reactor trips. Ann. Nucl. Energy 115, 280–293. https://doi.org/10.1016/j.anucene.2018.01.044
Pashutan, M., Abdolvand, N., Harandi, S.R., 2022. The impact of IT resources and strategic alignment on organizational performance: The moderating role of environmental uncertainty. Digit. Bus. 2, 100026. https://doi.org/10.1016/j.digbus.2022.100026
Patriarca, R., Di Gravio, G., Costantino, F., 2017. A Monte Carlo evolution of the Functional Resonance Analysis Method (FRAM) to assess performance variability in complex systems. Saf. Sci. 91, 49–60. https://doi.org/10.1016/j.ssci.2016.07.016
Patriarca, R., Falegnami, A., Costantino, F., Bilotta, F., 2018. Resilience engineering for socio-technical risk analysis: Application in neuro-surgery. Reliab. Eng. Syst. Saf. 180, 321–335. https://doi.org/10.1016/j.ress.2018.08.001

Patriarca, R., Ramos, M., Paltrinieri, N., Massaiu, S., Costantino, F., Di Gravio, G., Boring, R.L., 2020. Human reliability analysis: Exploring the intellectual structure of a research field. Reliab. Eng. Syst. Saf. 203, 107102. https://doi.org/10.1016/j.ress.2020.107102
Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, Inc., San Mateo, Calif.
Pedrosa, M.H., Salazar, A.K., Cardoso, C., Guedes, J.C., 2025. Study on Safety Culture Following the Implementation of a Near-Miss Management System in the Traditional Manufacturing Industry. Safety 11, 23. https://doi.org/10.3390/safety11010023
Pei, H., Gong, H., Bai, Z., Ma, Y., Xu, M., Li, G., 2024. A human factor reliability analysis method for maritime transport based on an improved CREAM model and group decision-making. Ocean Eng. 293, 116664. https://doi.org/10.1016/j.oceaneng.2024.116664
Peiravi, A., Nourelfath, M., Zanjani, M.K., 2022. Universal redundancy strategy for system reliability optimization. Reliab. Eng. Syst. Saf. 225, 108576. https://doi.org/10.1016/j.ress.2022.108576
Pence, J., Mohaghegh, Z., 2020. A Discourse on the Incorporation of Organizational Factors into Probabilistic Risk Assessment: Key Questions and Categorical Review. Risk Anal. 40, 1183–1211. https://doi.org/10.1111/risa.13468
Peng, C., Zhen, X., Huang, Y., 2023. Human-automation interaction centered approach based on FRAM for systemic safety analysis of dynamic positioning operations for offshore tandem offloading. Ocean Eng. 267, 113249. https://doi.org/10.1016/j.oceaneng.2022.113249
Provan, D.J., Woods, D.D., Dekker, S.W.A., Rae, A.J., 2020. Safety II professionals: How resilience engineering can transform safety practice. Reliab. Eng. Syst. Saf. 195, 106740. https://doi.org/10.1016/j.ress.2019.106740
Qiao, W., Liu, Y., Ma, X., Lan, H., 2021. Cognitive Gap and Correlation of Safety-I and Safety-II: A Case of Maritime Shipping Safety Management. Sustainability 13, 5509. https://doi.org/10.3390/su13105509
RAHAMAN, Md.A., UDDIN, Md.S., LATIF, W.B., 2023. Effects of Training and Development, Organizational Culture, Job Satisfaction, and Career Development on Employee Retention in Commercial Banks in Bangladesh. J. Asian Finance Econ. Bus. 10, 91–97. https://doi.org/10.13106/JAFEB.2023.VOL10.NO2.0091
Remenova, K., Skorkova, Z., Jankelova, N., 2018. Span of Control in Teamwork and Organization Structure. Montenegrin J. Econ. 14, 155–165. https://doi.org/10.14254/1800-5845/2018.14-2.10
Riemersma, B., Correljé, A.F., Künneke, R.W., 2024. Incorporating Safety-II in future gas systems. Saf. Sci. 173, 106462. https://doi.org/10.1016/j.ssci.2024.106462
Rosa, L.V., Haddad, A.N., De Carvalho, P.V.R., 2015. Assessing risk in sustainable construction using the Functional Resonance Analysis Method (FRAM). Cogn. Technol. Work 17, 559–573. https://doi.org/10.1007/s10111-015-0337-z
Salmon, P., Jenkins, D., Stanton, N., Walker, G., 2010. Hierarchical task analysis vs. cognitive work analysis: comparison of theory, methodology and contribution to system design. Theor. Issues Ergon. Sci. 11, 504–531. https://doi.org/10.1080/14639220903165169
Shi, H., Wang, J.-H., Zhang, L., Liu, H.-C., 2023. New improved CREAM model for human reliability analysis using a linguistic D number-based hybrid decision making approach. Eng. Appl. Artif. Intell. 120, 105896. https://doi.org/10.1016/j.engappai.2023.105896
Slim, H., Nadeau, S., 2020. A Mixed Rough Sets/Fuzzy Logic Approach for Modelling Systemic Performance Variability with FRAM. Sustainability 12, 1918. https://doi.org/10.3390/su12051918
Srinivas, S., 1993. A generalization of the noisy-or model. In: Proceedings of the Ninth international conference on Uncertainty in artificial intelligence. Elsevier, pp. 208-215.
Stanton, N.A., 2006. Hierarchical task analysis: Developments, applications, and extensions. Appl. Ergon. 37, 55–79. https://doi.org/10.1016/j.apergo.2005.06.003
Stanton, N.A., 2004. Systematic human error reduction and prediction approach (SHERPA). In Handbook of human factors and ergonomics methods (pp. 394-403). CRC Press.
Sun, H., Wang, H., Yang, M., Reniers, G., 2022. A STAMP-based approach to quantitative resilience assessment of chemical process systems. Reliab. Eng. Syst. Saf. 222, 108397. https://doi.org/10.1016/j.ress.2022.108397

Sun, Z., Li, Z., Gong, E., Xie, H., 2012. Estimating Human Error Probability using a modified CREAM. Reliab. Eng. Syst. Saf. 100, 28–32. https://doi.org/10.1016/j.ress.2011.12.017

Tang, Y., Zhang, X., Zhou, Y., Huang, Y., Zhou, D., 2023. A new correlation belief function in Dempster-Shafer evidence theory and its application in classification. Sci. Rep. 13, 7609. https://doi.org/10.1038/s41598-023-34577-y

Wahl, A., Kongsvik, T., Antonsen, S., 2020. Balancing Safety I and Safety II: Learning to manage performance variability at sea using simulator-based training. Reliab. Eng. Syst. Saf. 195, 106698. https://doi.org/10.1016/j.ress.2019.106698

Wang, F., Tian, J., Lin, Z., 2020. Empirical study of gap and correlation between philosophies Safety-I and Safety-II: A case of Beijing taxi service system. Appl. Ergon. 82, 102952. https://doi.org/10.1016/j.apergo.2019.102952

Wang, Y., Li, P., Hong, C., Yang, Z., 2025. Causation analysis of ship collisions using a TM-FRAM model. Reliab. Eng. Syst. Saf. 260, 111035. https://doi.org/10.1016/j.ress.2025.111035

Wang, Y., Zhang, S., Xu, S., 2022. Impact of Efficient Resource Management Practices on Sustainable Performance: Moderating Role of Innovative Culture-Evidence From Oil and Gas Firms. Front. Psychol. 13, 938247. https://doi.org/10.3389/fpsyg.2022.938247

West, J., Siddhpura, M., Evangelista, A., Haddad, A., 2024. Improving Equipment Maintenance—Switching from Corrective to Preventative Maintenance Strategies. Buildings 14, 3581. https://doi.org/10.3390/buildings14113581

Wu, W., Cantero-Chinchilla, S., Prescott, D., Remenyte-Prescott, R., Chiachío, M., 2024. A general approach to assessing SHM reliability considering sensor failures based on information theory. Reliab. Eng. Syst. Saf. 250, 110267. https://doi.org/10.1016/j.ress.2024.110267

Xi, Y.T., Yang, Z.L., Fang, Q.G., Chen, W.J., Wang, J., 2017. A new hybrid approach to human error probability quantification–applications in maritime operations. Ocean Eng. 138, 45–54. https://doi.org/10.1016/j.oceaneng.2017.04.018

Xie, S., Huang, Z., Wu, G., Luo, J., Li, L., Ma, W., Wang, B., 2024. Combining precursor and Cloud Leaky noisy-OR logic gate Bayesian network for dynamic probability analysis of major accidents in the oil depots. Reliab. Eng. Syst. Saf. 241, 109625. https://doi.org/10.1016/j.ress.2023.109625

Xu, D.-L., Yang, J.-B., 2005. Intelligent decision system based on the evidential reasoning approach and its applications. J. Telecommun. Inf. Technol. 73–80. https://doi.org/10.26636/jtit.2005.3.323

Xue, J., Yang, P., Wang, Z., Li, Q., Hu, H., 2025. Cause analysis and management strategies for ship accidents: A Bayesian decision support model. Ocean Eng. 320, 120291. https://doi.org/10.1016/j.oceaneng.2025.120291

Yang, Q., Tian, J., Zhao, T., 2017. Safety is an emergent property: Illustrating functional resonance in Air Traffic Management with formal verification. Saf. Sci. 93, 162–177. https://doi.org/10.1016/j.ssci.2016.12.006

Yang, X., Zhou, T., Utne, I.B., Zhang, W., 2025. A framework for quantification of coupling risk in the transition between operational modes of MASS. Reliab. Eng. Syst. Saf. 264, 111436. https://doi.org/10.1016/j.ress.2025.111436

Yang, Z., Lau, Y., Poo, M.C.-P., Yin, J., Yang, Zaili, 2025. Port resilience to climate change in the Greater Bay Area. Transp. Res. Part Transp. Environ. 142, 104681. https://doi.org/10.1016/j.trd.2025.104681

Yang, Z.L., Bonsall, S., Wall, A., Wang, J., Usman, M., 2013. A modified CREAM to human reliability quantification in marine engineering. Ocean Eng. 58, 293–303. https://doi.org/10.1016/j.oceaneng.2012.11.003

Yin, J., Khan, R.U., Wang, X., Asad, M., 2024. A data-centered multi-factor seaport disruption risk assessment using Bayesian networks. Ocean Eng. 308, 118338. https://doi.org/10.1016/j.oceaneng.2024.118338

Yu, M., Quddus, N., Kravaris, C., Mannan, M.S., 2020. Development of a FRAM-based framework to identify hazards in a complex system. J. Loss Prev. Process Ind. 63, 103994. https://doi.org/10.1016/j.jlp.2019.103994

Yu, Y., Liu, K., Fu, S., Chen, J., 2024. Framework for process risk analysis of maritime accidents based on resilience theory: A case study of grounding accidents in Arctic waters. Reliab. Eng. Syst. Saf. 249, 110202. https://doi.org/10.1016/j.ress.2024.110202

Zarei, E., Khan, F., Abbassi, R., 2022. A dynamic human-factor risk model to analyze safety in sociotechnical systems. Process Saf. Environ. Prot. 164, 479–498. https://doi.org/10.1016/j.psep.2022.06.040

Zhao, Y., Ma, X., Qiao, W., Zhang, J., 2025. Resilient maritime transportation system from the perspective of FRAM: conceptualization and assessment. Reliab. Eng. Syst. Saf. 262, 111155. https://doi.org/10.1016/j.ress.2025.111155

Zhou, Q., Wong, Y.D., Loh, H.S., Yuen, K.F., 2018. A fuzzy and Bayesian network CREAM model for human reliability analysis – The case of tanker shipping. Saf. Sci. 105, 149–157. https://doi.org/10.1016/j.ssci.2018.02.011

Zinetullina, A., Yang, M., Khakzad, N., Golman, B., Li, X., 2021. Quantitative resilience assessment of chemical process systems using functional resonance analysis method and Dynamic Bayesian network. Reliab. Eng. Syst. Saf. 205, 107232. https://doi.org/10.1016/j.ress.2020.107232

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Zaili Yang reports financial support was provided by European Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.