

An Operational Risk Analysis Model for Container Shipping Systems considering Uncertainty Quantification

Reliability Engineering and System Safety

Son Nguyen ^{a,b}, Peggy Shu-Ling Chen ^a, Yuquan Du ^{a*}, Vinh V. Thai ^c

^a National Centre for Ports and Shipping, Australian Maritime College, University of Tasmania, TAS 7248, Australia

^b Department of Maritime Transportation Economics, Faculty of Economics, Vietnam Maritime University, Hai Phong, Vietnam

^c School of Business IT & Logistics, RMIT University, Melbourne, VIC 3000, Australia

*: Corresponding author

Emails:

Son Nguyen: son.nguyen@utas.edu.au

Peggy Shu-Ling Chen: p.chen@utas.edu.au

Yuquan Du*: yuquan.du@utas.edu.au (Phone: +61 3 632 431 23)

Vinh V. Thai: vinh.thai@rmit.edu.au

ORCID

Son Nguyen <https://orcid.org/0000-0001-8736-8454>

Peggy Shu-Ling Chen: <https://orcid.org/0000-0002-1513-4365>

Yuquan Du: <https://orcid.org/0000-0001-7540-493X>

Vinh V. Thai: <https://orcid.org/0000-0001-6863-8439>

An Operational Risk Analysis Model for Container Shipping Systems considering Uncertainty Quantification

ABSTRACT

Different uncertain factors obstruct the analysis of operational risks in container shipping, especially those rooted from the subjectivity of multiple risk assessments and their aggregation. This paper proposes a risk analysis model featuring a quantification of the uncertainty level for each risk. Subjective probabilities (degree of belief) of a risk at different states is quantified by a Bayesian network based on a two-level parameter set. A set of three uncertainty indicators are developed to provide a quantitative diagnosis of the knowledge base, including *expert ignorance*, *disagreement among experts*, and *polarization of their assessments*. The relative risk situation is then visualized by risk mapping, using both Risk index and Uncertainty index. Investigating uncertainty through both subjective and objective indicators gains insights into the reliability of risk analysis and assessment. The empirical experiment is carried out on a shipping company not only confirmed the feasibility of the proposed model but also demonstrated its usefulness in analyzing highly uncertain risks.

Keywords: Risk assessment; Container shipping operation; Bayesian Network; Evidential Reasoning; Risk mapping

1. INTRODUCTION

The fast growth in complexity and scale puts container shipping systems under various risks posed by potential unfortunate events that occur in a company's daily operations (Chang et al., 2015). Operational risks are considered as an important risk sector by stakeholders with the highest likelihood and pervasive impacts (Lam and Lassa, 2016, Liu et al., 2018). Undetected operational threats such as cargo misdeclaration, network infiltration, piracy and unexpected financial disruptions can lead to significant consequences for shipping companies. The hazardous events (HEs) such as the NotPetya attack, Tianjin port explosion, Maersk Honam fire, and Hanjin bankruptcy are well-observed by the industry. Container shipping operational risk (CSOR) management requires risk mitigation/prevention (RMP) plans that depend heavily on dedicated but limited resources (Chang et al., 2014, Lee and Song, 2017). Effective, informative, and holistic risk analysis, therefore, is crucial for a corroborated and confident decision-making process.

It is agreed in risk foundational studies that uncertainty should be comprehensively considered for an adequate and systematic risk approach (Levin, 2005, Aven and Zio, 2011). However, this objective is seemingly problematic for risk assessment in container shipping. On top of the system complexity, the physical movements of shipments require synchronization from both information and financial flows to maintain its seamlessness. The reliability of the system, therefore, is heavily affected by the accumulation of uncertain factors from its components (e.g., communication infrastructure, vehicles, human operations, handling equipment, containers, and shipments) (Ottomanelli and Wong, 2011). Although uncertainties are all epistemic (i.e., rooted in the lack of knowledge), their sources can be used to classify them. The non-determinate nature of a future HE causes *outcome uncertainty* (U_O). In the container shipping context, it is observable that events of a same risk are hardly the same and even greatly different from one another. To narrow-down this variation in the CSOR analysis, expert assessments were widely implemented to indicate the most likely scenarios (see, for example, Chang et al. (2015)). However, the weakness of the knowledge base used for these assessments poses another type of uncertainty, termed by Levin (2005) as *evidential uncertainty* (U_E). While most quantitative risk analysis (QRA) models handle uncertainty only through probability, their ability to investigate and communicate both U_O and U_E was scarcely featured (Goerlandt and Montewka, 2015b).

This study develops a novel QRA model for CSORs that has the ability to quantitatively detect the uncertainty level in the process of risk assessment. The core of our solution lies in the set of three uncertainty indicators. *Expert's ignorance*, which conveys the subjective assessments of U_E , is expressible through the concept of Degree of Belief (DoB) and Unassigned DoB (UDoB). Since this attribute might be vulnerable against subjective biases, the *average disagreement among experts* is employed to provide a more objective view of U_E . *Assessment polarization* captures the U_O by the degree to which a scenario cannot be indicated as more likely to happen than the others. This

investigation is able to reveal cases of problematic implemented knowledge base. The enabling of uncertainty quantification allows the level of risk to be evaluated while maintaining an informative description of the knowledge base that render the reliability of the analysis results.

The remaining of this paper is structured as follows. Section 2 presents an analytical review to describe the knowledge gap as well as develop the key concepts of the QRA model. Section 3 proposes the QRA model composed of three main components: (1) The Evidential Reasoning (ER) algorithm to aggregate multiple expert assessments, (2) a Bayesian Network (BN) model that calculates the level of CSORs and (3) an uncertainty quantification module (UQM) based on uncertainty indicators that gauges the credibility of the CSOR analysis. The model validation process, including an illustrative case study, a validity assessment and a sensitivity analysis is introduced in Section 4. Finally, discussions and conclusions are provided in Section 5.

2. LITERATURE REVIEW AND CONCEPT DEVELOPMENT

Through reviewing and analyzing relevant literature, this section articulates an in-depth research background analysis with methodological-focused discussions to formulate the key concepts of the QRA. The contributions to literature of this study are presented at the end of the section.

2.1. The need for an assessment model with effective uncertainty treatment for CSORs

The current risk understanding considers *The prior HE* (A), its *Consequence* (C), and the involved *Uncertainty* (U) as the three main elements in the concept of risk ($R = A, C, U$) (Zio, 2007, Aven, 2012). Winkler (1996) concluded that all uncertainties are essentially epistemic. However, we concur with Aven and Renn (2010) that an effective differentiation is handy in capturing uncertainty comprehensively as well as drawing useful insights into the risk situation. Since the HE is non-determinate in the present timeframe, U_O is inherently irreducible (Levin, 2005, Aven, 2012). On the other hand, U_E is reducible through improving the quality of the knowledge base (Apeland et al., 2002). The differentiation and dedicated treatments of U_O and U_E is the key to complete awareness and proactive handling of U .

Although the incremental role of risk analysis in maritime transport has been shown in subsections such as navigation safety (Goerlandt and Montewka, 2015b), and offshore human elements (Eleye-Datubo et al., 2008, Akhtar and Utne, 2014), the attention of the maritime academia was superficially paid in the risk managerial perspective, particularly with CSORs (Lee and Song, 2017). In addition, studies in maritime transport risk suffer from lack of uncertainty awareness and anticipation (Goerlandt and Montewka, 2015b). Studies often consider probability as the core of the risk concept while uncertainty was not communicated any further than its probabilistic description. For example, in the CSOR analysis conducted by Chang et al. (2014), Chang et al. (2015) as well as those related to maritime shipping systems by Berle et al. (2011), and Yang (2011) the subjective assessments could be vague and

uncertain, and thus difficult for a crisp scale without specific descriptions to capture effectively. Using linguistic assessments in combination with fuzzy theory can provide a partial expression of U_E (see, for example, Tseng et al. (2013)). However, this approach lacks a distinguishable realization of U_O and U_E . The assessors could not express their extents of confidence, which heavily depend on the base of knowledge, in their assessments. An uncertainty-aware QRA model is necessary to promote more reliable practices in the maritime risk research community, considering that not yet a concrete cost-effectiveness validation method for QRA is available (Aven and Heide, 2009, Goerlandt et al., 2017).

2.2. Subjective risk assessments and the Uncertainty Quantification Module (UQM)

A realist risk approach is unlikely to be compatible with assessing multiple CSORs. *Ex post* risk analyses require enriched risk historical databases as the primary input for risk assessment. A sizeable and compatible recorded risk database, however, is difficult to gather in the field of container shipping. Additionally, an anticipative QRA result is needed for timely RMP (Aven and Krohn, 2014). Relying exclusively on historical data effectively assumed that an HE can be repeated multiple times to establish an objective probability (the law of large number), which cannot be justified in a highly dynamic risk situations as container shipping operations. A promising alternative here is a rational constructivist approach that involves expert predictive assessments. The unavoidable subjective biases are controlled by the connections between these assessments and reality through a knowledge base of evidence and reasonings (Rae and Alexander, 2017). The insights into the quality of the implemented knowledge base, therefore, crucial in constituting the reliability of the QRA.

A set of DoBs and UDoB can be utilized as a form of subjective probability that reflect the experts' distributions of likelihood for exclusive states of the risk parameters (e.g., Financial impact: 50% Low, 20% Medium, 10% High, and 20% UDoB). The *average disagreement among experts* is computed in this stage (Figure 1, Indicator 1). UDoB is a means to subjectively express the weakness of the knowledge base. Here, ER, rooted in the Dempster – Shafer theory (Shafer, 1976), can be used to aggregate multiple experts' assessment of a risk parameter and extract aggregated UDoBs to have the *total expert's ignorance* indicator (Figure 1, Indicator 2). The *total polarization* of normalized aggregated DoBs, which is now contained merely risk magnitude information, can also be calculated (Figure 1, Indicator 3). In this system, Indicator 1 and 2 indicate the weakness of the knowledge base related to the available evidence (U_E) while Indicator 3 captures the outcome uncertainty degree to which a scenario cannot be identified as more likely to happen than others (U_O).

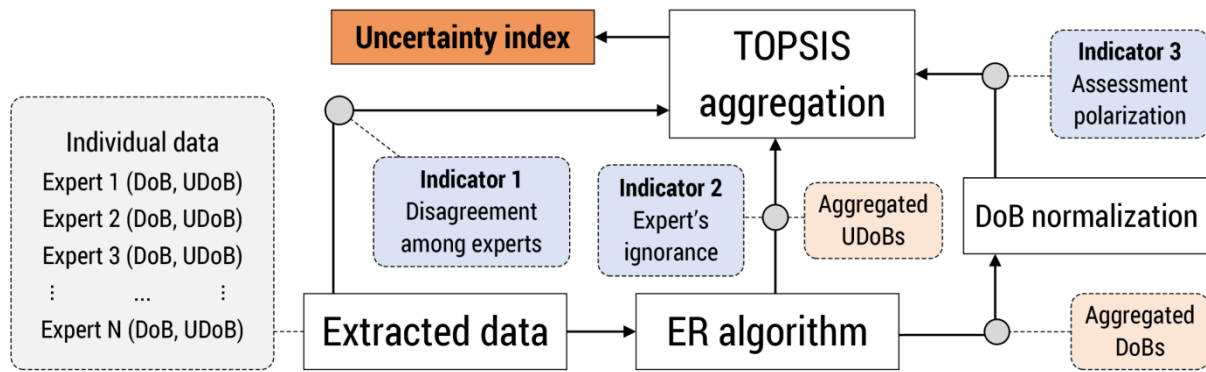


FIGURE 1 Design of the proposed Uncertainty Quantification Module (UQM)

The aggregation of these indicators is carried out by the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), which was proposed by Hwang and Yoon (1981). TOPSIS is suitable to aggregating multiple numeric values of incongruous dimensions and monotonically varying. In our model, TOPSIS is employed to aggregate three uncertainty indicators to derive the Uncertainty index (UI) for each risk, which provides insights into the foundation of the risk assessments. The integration of DoB, UDoB, ER, and TOPSIS forms the UQM of the QRA model. Its mathematical description will be presented in Section 3.3.

2.3. The integration of BN into the risk parameter structure

BN is well known for its implementations in risk assessment and analysis practices (Aven, 2012, Duijm, 2015). BNs allow correlation among variables and predictions to be made even when direct evidence or observations are missing (Krieg, 2001). Integrating the Bayesian theorem into risk parameter set is a technique in quantitative assessment of CSORs (see, for example, the QRA models of Yang et al. (2008) and Alyami et al. (2014) with the parameters based on Failure Mode and Effects Analysis (FMEA)). The BN model proposed by Yang et al. (2008) is able of calculating failure priority value at the cost of bulky and arduous quantification of all causal relationships. Alyami et al. (2014) proposed a simpler but more rigid method where the DoB in the states are distributed linearly with the evidence in risk parameters. The importance of parameters is expressible in the model of Wan et al. (2019). However, uncertainty was not adequately considered even though cases of notable difference between experts' assessments are observable in the applications of both Alyami et al. (2014) and Wan et al. (2019)'s models. For instance, averaging 0% and 90% assessments to 45% effectively ignored crucial phenomena of the knowledge base (e.g., ambiguity, conflicts of understandings). Again, an UQM developed for this family of models is significant considering these potential weaknesses of the risk analysis.

Additionally, there is a gap from extant literature of a rational-structured parameter set for CSOR. *Firstly*, the tailored-made parameter sets (e.g., the four-parameter model proposed by Alyami et al.

(2014)) increased the fragmentation of risk parameter sets in the literature. This multiplicity, however, obstructs the communications of QRA results among CSOR studies and, therefore, impedes the establishment of a robust CSOR understandings (Goerlandt and Montewka, 2015b, Aven and Zio, 2014). *Secondly*, the generality and, in some cases, even vagueness in the meaning of parameters might cause ambiguity in the input extraction process and affect the reliability of the QRA. For example, the *impact to the resilience of port* in the model of Alyami et al. (2014) can be understood as a form of consequence though it was structured as a separate risk parameter; the parameter of *damage to quality* in the model of Wan et al. (2019) is difficult to be distinguished from the physical damage to shipments and infrastructures, and the quality of the transport services, which is expressible in term of *delays (time)*. Therefore, a clear-cut parameter set based on the perspective of bearers in the context of CSOR is crucial to develop a reliable QRA model.

2.4. An interpretative approach to risk mapping

Risk mapping is a widespread tool for analyzing risks (Ale et al., 2015, Goerlandt and Reniers, 2016, Cox, 2008). In CSOR studies, risk mapping has been used extensively in risk prioritization and RMP strategies (see, for example, the applications of Yang (2011), Tseng et al. (2013), and Chang et al. (2014)). However, multiple disadvantages of this method were mentioned in the literature. *Firstly*, the knowledge base was not communicated in these risk maps, uncertainty is only visualized through probability or likelihood. *Secondly*, Cox (2008) proved that the risk of matrices could be worse than useless if the matrix is poorly constructed (i.e., size, system of classification), especially if U_O of risks are high. *Thirdly*, “risk ties”, mainly due to the low resolution of risk maps (Goerlandt and Reniers, 2016, Duijm, 2015, Cox, 2008), exist in the risk maps of Yang (2011), Tseng et al. (2013) and Chang et al. (2014).

Several proposals were recommended to address these flaws such as design normative properties (Cox, 2008), uncertainty presence (Goerlandt and Reniers, 2016), continuous diagram (Duijm, 2015), and three-dimensional presentation (Aven, 2013). Yet, although more data could be jam-packed into a risk diagram, the limitation of both typesetting options (Goerlandt and Reniers, 2016) and the ability of human brain in perceiving effectively more than three spatial dimensions seem to hinder the application of risk mapping as a primary risk analysis tool. In this study, risk mapping is implemented as a tool to support demonstrating overall risk situation, including both quantified risk levels and the attached uncertainty. This application exploits the forte of risk mapping as a straightforward, decision-making support and conspicuous method for risk visualization and interpretation.

2.5. Research objectives, novelty, and contributions to existing literature

To fill the gap of a managerial QRA model for CSORs that is capable of detecting the weakness of the knowledge base, this study (1) recognizes the importance of uncertainty awareness in the process of

risk quantification; (2) designs a module to gain insights into the uncertainty situation of risk and address it beyond the traditional probability expression; and (3) enables the expression and consideration of both U_O and U_E throughout the QRA. This paper has three primary novelties as presented below.

- **This study confirms the criticality of uncertainty awareness in risk assessment.** Through the analysis of literature and the case study, uncertainty recognition and the separation of U_O and U_E are proved critical to the reliability of QRA. The realization of uncertainty and its visualization by risk mapping opened another dimension of RMP strategy in continuously improving the knowledge base.

- **A sound QRA model for CSOR assessment and prioritization is developed considering both U_O and U_E .** The combination of multiple methodologies including BN, ER, and TOPSIS facilitate a systematic capturing of uncertainty in assessing CSORs to support decision-making. The four qualities of uncertainty handling for QRAs considered in this paper are as follows:

1. The concept of R is introduced unambiguously.
2. U is considered as a main component of, but not R .
3. U_O and U_E are differentiated from one another.
4. U_O and U_E are expressible in the quantification process.

- **The case study validates two important motivations of this study.** *Initially*, the case study pragmatically confirms the phenomenon of weak knowledge base in quantitatively analyze CSORs based on subjective probability distributions. *Additionally*, the integration of the UQM reveals the instances of experts' limited ability in assessing the knowledge base of the whole panel.

This study has also made methodological contributions including:

- **To CSOR assessment and analysis studies:** (1) *A well-defined two-level BN is established for CSOR assessment.* Parameters are supported by events experienced by the industry and presented with clear and customizable definitions of states. (2) *ER is employed to aggregate and separate UDoBs from DoBs.* The ER algorithm allows the subjective expression of U_E and the separation of U_E and U_O . This ability enables the rational calculation of assessment polarization (capturing U_O), which was previously obstructed by the intertwining of U_E and U_O in each risk assessment.

- **To Bayesian Network based risk assessment studies:** *An UQM with a set of one subjective and two objective indicators is integrated into QRA.* The UQM allows an insightful and informative quantification of uncertainties while preserving the intersubjectivity of QRA inputs. Based on the results from the UQM and risk map, solutions to improve the knowledge base for risk assessment could be inferred.

- **To the application of risk mapping:** *Risk mapping is not used directly to assess risks but visualize the situation of risk and the attached weakness of the knowledge base.* Risk level (Risk index) is displayed in combination with Uncertainty level (Uncertainty index) to depict an interpretative overview that supports risk-related decision-making.

3. AN ADVANCED CSOR ASSESSMENT MODEL WITH UNCERTAINTY EVALUATION

This section introduces the QRA model. The input data are extracted from the expert panel and then processed by ER to separate the DoBs (risk level data) from the attached UDoBs (uncertainty data). Three uncertainty indicators are calculated along the process of data manipulation by the UQM and put into TOPSIS to calculate Uncertainty Index (UI). The normalized aggregated DoBs are fed into the BN model to derive Risk Index (RI), which will later be presented together with UIs to describe the risk situation in a risk map. A stepwise illustration of the model is in Figure 2.

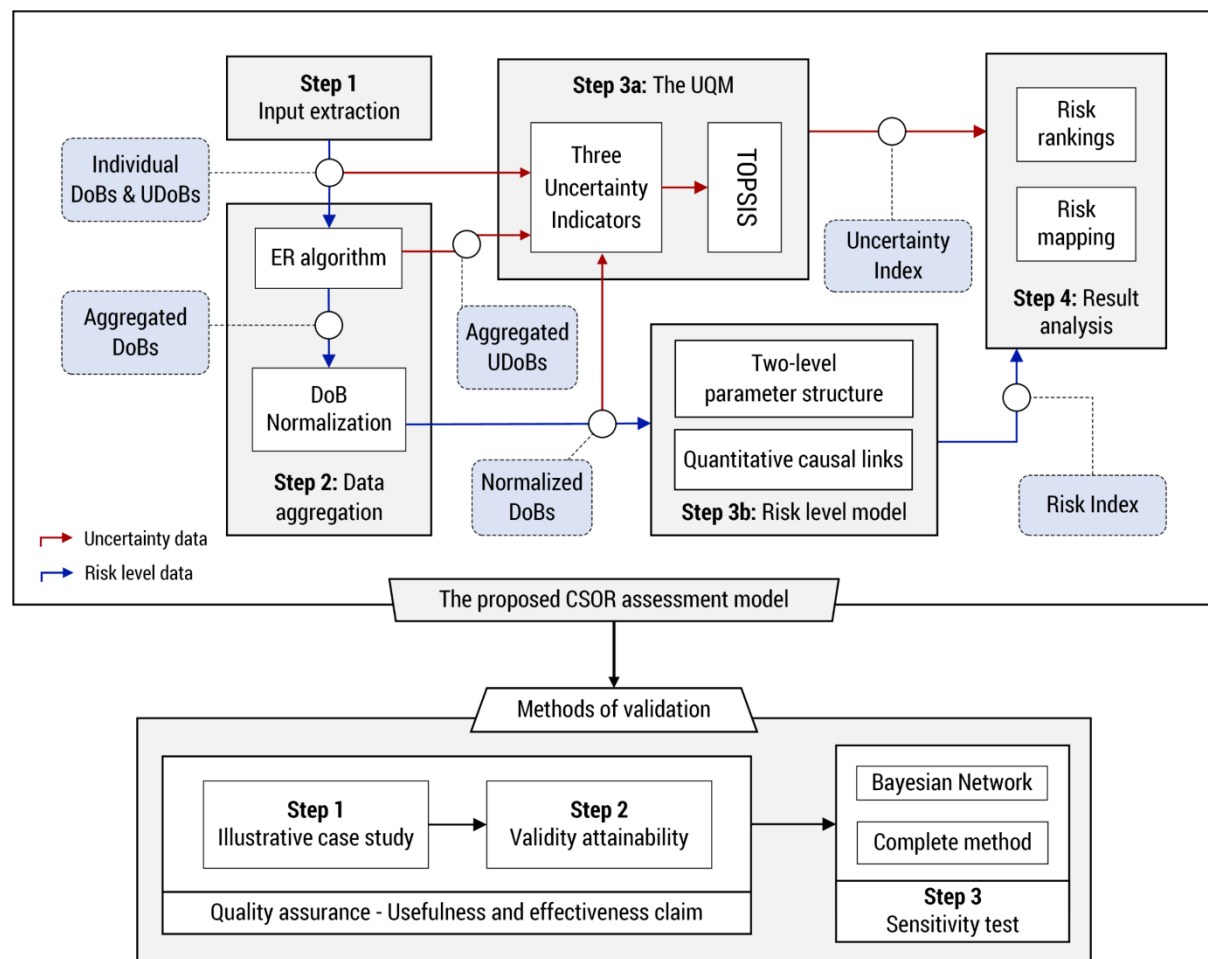


FIGURE 2 The structure of CSOR assessment model and methods of validation

3.1. Input extraction

Risk assessments are first extracted from multiple experts in the form of subjective probability distributions (see Section 2.2 for an example). Each assessment includes (1) DoBs of that expert for different states of a parameter, and (2) a UDoB, so-called “ignored belief mass” that expresses the ignorance/lack of knowledge in that individual assessment (Yang and Xu, 2002). In other words, UDoBs represent the weakness of the knowledge base (U_E) in conjecturing the DoBs in different states. For each assessment task, experts will first deduct an amount of UDoB from a whole unit of belief (100%) according to their judgment on the quality of the acquired knowledge base (i.e., understandings of the related phenomena, and the availability and reliability of the collected evidence). Then, the remainder can be distributed to the states of the assessed subject as DoBs. Other available supplementary sources of data such as recorded historical events, operational reports, forecasts, and third-party consultations should also be utilized to support experts in this step.

3.2. Risk assessments aggregation by the ER algorithm

In this step, DoBs and UDoBs from experts will be mathematically aggregated by the ER algorithm. ER is developed based on the evidence combination rule of the Dempster–Shafer theory that allows the aggregation of multiple sets of beliefs. Unlike the arithmetic average method, a set of probabilities based on a significant foundation of evidence and understandings should not be considered as equal with those that are not. The degree to which the weakness of the knowledge base is presented in ER in the form of “ignored belief mass” – the amount of probability that is not distributed to any possibility. Those characteristics of ER are presented through an example illustrated in Table 1. This study uses the ER algorithm proposed by Yang and Xu (2002) for its superiority in satisfying synthesis axioms.

TABLE 1 An example of combining subjective probability using arithmetic average and ER

	Expert 1	Expert 2	Arithmetic average	ER algorithm
Possibility 1	0.5	0.05	0.275	0.3218
Possibility 2	0.2	0.1	0.15	0.1673
Possibility 3	0.1	0.25	0.175	0.1764
Ignored belief mass	0.2	0.6	0.4	0.3345

In our model, risk knowledge is divided into two disparate realms: the risk evaluation data (DoBs) – the belief mass, and the ignorance realization (UDoBs) – the ignored belief mass. Assume that there are N experts and sets of probabilities regarding M states of an individual risk parameter. Let d_{mn} be the DoB that the expert n assigned for the state m and u_n be the remained unassigned belief mass of the expert n . The results include aggregated DoBs (D_1, \dots, D_M) and aggregated UDoB DU , which is an expression of U_E (Equation 1). It is noteworthy that ER also supports the difference in importance of

the experts, but since there are no reliable method and evidence to quantify that aspect, their assessments are considered as equally important in this study. DoBs are then normalized using Equation 2.

$$ER \begin{pmatrix} d_{11} & \cdots & d_{N1} \\ \vdots & \ddots & \vdots \\ d_{M1} & \cdots & d_{MN} \\ u_1 & \cdots & u_N \end{pmatrix} = \begin{bmatrix} D_1 \\ \vdots \\ D_M \\ DU \end{bmatrix} \quad (1)$$

$$\hat{D}_m = \frac{D_m}{\sum_{m=1}^M D_m} \quad m = 1, \dots, M \quad (2)$$

3.3. Evaluation of the uncertainty situation – The UQM

The uncertainty in CSOR assessment will be quantified in a relative manner since it is difficult and unnecessary to establish a baseline for such concept. Qualitative classification schemes for this purpose could be found in the studies of Goerlandt and Montewka (2015a) and Askeland et al. (2017). While U_E is expressible through the concept of UDoB, it is subjective and only based on the individual knowledge base, thus cannot depict the whole picture of uncertainty. Hence, this study considers three factors as the main indicators for the strength of knowledge: (1) *Average disagreement among experts*; (2) *Expert's ignorance*, expressed by aggregated UDoBs; and (3) *Expert assessment polarization* (see Section 2.2). All indicators are weighted in the calculation process with each parameter (ω_c) and then normalized (denoted with a “^” symbol) for a set of Y risks.

(1) Average disagreement among experts (Δ_A): The average discrepancy among assessments provided by N experts across all M states and C parameters. Since DoBs and UDoBs are not discriminated in this indicator, it is computed directly from the extracted assessments from step 1 (Equation 3). In the dividend, d_{cmj} and d_{cmk} are the DoBs of expert j and k on parameter c and state m .

$$\Delta_A = \frac{\sum_{c=1}^C \sum_{m=1}^M \sum_{\substack{j=1 \\ j < k}}^N \omega_c |d_{cmj} - d_{cmk}|}{\frac{N(N-1)}{2}} \quad \hat{\Delta}_{A_i} = \frac{\Delta_{A_i}}{\sqrt{\sum_{i=1}^Y \Delta_{A_i}^2}} \quad (3)$$

(2) Total expert ignorance expressed by UDoBs ($\sum DU$): $\sum DU$ of each risk is the weighted sum of UDoBs, which are denoted as DU_c with parameter c . $\sum DU$ is calculated using Equation 4 with results derived after the ER aggregation step (Equation 1).

$$\sum^{DU} = \sum_{c=1}^C \omega_c DU_c \quad \widehat{\sum^{DU}}_i = \frac{\sum^{DU}_i}{\sqrt{\sum_{i=1}^Y \sum^{DU}_i^2}} \quad (4)$$

(3) Total expert assessment polarization $\sum \Delta_P$: The deviation of the expert assessment from the totally random assessment, denoted by ξ with $\xi = M^{-1}$. It is computed using Equation 5 based on the normalized aggregated DoBs \widehat{D}_m derived after ER (Equation 2).

$$\sum \Delta_P = \sum_{c=1}^C \omega_c \sum_{m=1}^M |\widehat{D}_m - \xi| \quad \widehat{\sum \Delta_P}_i = \frac{\sum \Delta_P_i}{\sqrt{\sum_{i=1}^Y \sum \Delta_P_i^2}} \quad (5)$$

TOPSIS is employed here to aggregate the uncertainty indicators. For each risk, we have three data entries $R_i = (\widehat{\Delta}_A, \widehat{\sum DU}_i, \widehat{\sum \Delta_P}_i)$ ($i = 1, 2, \dots, Y$). The geometric distances from each risk to the positive and negative “ideal risks” (PIR and NIR in Equation 6) that have the superlative uncertainty conditions is used to calculate UI (Equation 7). It is noteworthy that unlike Δ_A and $\sum DU$, $\sum \Delta_P$ negatively correlates with the overall uncertainty level. Higher UI indicates a higher level of uncertainty and lower strength of knowledge base.

$$\begin{aligned} \text{PIR} &= \{\max(\widehat{\Delta}_A), \max(\widehat{\sum DU}), \min(\widehat{\sum \Delta_P})\} \\ \text{NIR} &= \{\min(\widehat{\Delta}_A), \min(\widehat{\sum DU}), \max(\widehat{\sum \Delta_P})\} \end{aligned} \quad (6)$$

$$UI_i = \frac{\sqrt{\sum (R_i - \text{NIR})^2}}{\sqrt{\sum (R_i - \text{PIR})^2} + \sqrt{\sum (R_i - \text{NIR})^2}} \times 100\%, (i = 1, 2, \dots, Y) \quad (7)$$

3.4. Risk level assessment model

A BN model is developed to calculate the relative risk level. Two primary components of a BN are *the network structure* and *the quantitative causal relationships*. A system of three states (Low, Medium, High) and the reasoning mechanism of BN are implemented to cover outcome uncertainty (U_O). The Risk Index (RI) values are calculated to prioritize risks.

3.4.1. A two-level parameter set for CSOR assessment

The traditional approach of FMEA uses only three parameters (*Likelihood of occurrence* (L); *Severity of consequence* (S) and *Probability of being undetected* (U)) to measure *Risk level* (R). However, the excessive generality of the parameters, especially S , might affect the ability of experts in giving their assessments. Our model implements the parameter set proposed by Nguyen and Wang (2018) with

explicit definition and example for each parameter (Figure 3).

From the view of a container shipping company, the consequences of an HE could be separated into three aspects: *Financial impact (F)*, *Reputational impact (I)* and *Operational impact (O)*. **F** captures the damage on the revenue of the company and is expressed through financial losses (*e.g.*, fines, additional fees, damages to infrastructures, or compensations). **I** captures the negative effects of the event toward the company's credibility in the view of customers as well as their standing in the industry perceived by partners (Stopford, 2009, Yuen et al., 2018), which could be observed through complaints, breach of contracts or agreement, or decreases in stock prices. **O** captures the effects on the operational plans of a shipping company including unexpected but inescapable adjustments in voyage schedule and reallocation or intensification of human resources in response to the consequences (Bakshi et al., 2011, Nguyen and Wang, 2018). An example is the NotPetya cyberattack on A.P. Møller-Maersk in 2017. The attack was described in detail by Greenberg (2018). Maersk estimated a 250-300 million USD loss, which can be considered as **F**. **I** is observable through the fact that many partners of the company had to abruptly change their transport plan and even production with much higher costs; new customers had to book their slot through inconvenient channels of information. **O** of this attack is relatively heavy. Normal operations of the company was totally disrupted and human resources had to be rearranged to respond to the HE (*e.g.*, handle new booking orders, reestablish the IT system including computers and servers).

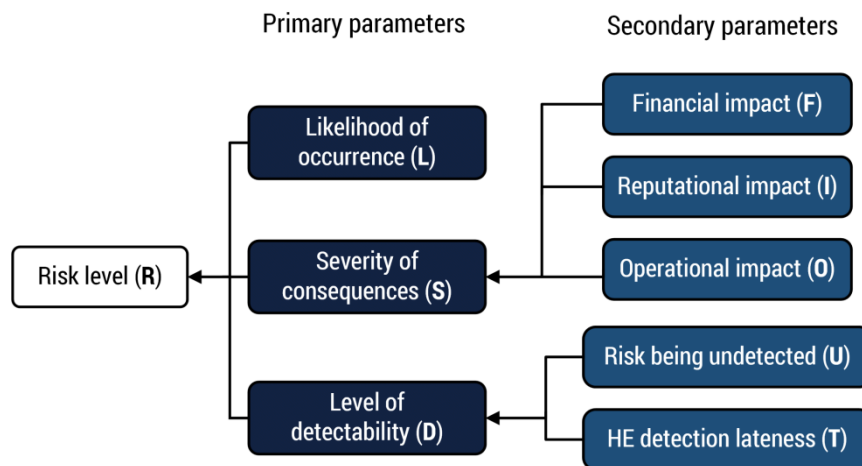


FIGURE 3 The risk parameter structure for CSOR (Nguyen and Wang, 2018)

The ability to detect instances of risk and its HE is presented by the Level of detectability (*D*). In container shipping, a more “hidden” and “self-concealed” HE is emphatically more dangerous since it is more difficult to be detected and hence, harder to be prevented by risk-bearers. For example, in the case of Maersk Honam, the risk of fire originated from flammable cargoes was already well-known, but it is still relatively hard to be early detected and prevented because of cargo misdeclaration (Porter, 2018). Risk being undetected (*U*) here is the parameter by which such a characteristic could be assessed.

Besides, the HE detection lateness (T) is also critical since it negatively affects the effectiveness of companies' responses to limit the damages caused by the occurrence of a particular HE. For instance, an outage of a reefer stack is usually detected and resolved within several hours (Filina and Filin, 2008) while a cargo self-ignition is harder to be noticed, eliminated, or effectively controlled within minutes before it spreads to other containers or causing explosions.

3.4.2. A BN model for CSOR assessment

Normalized DoBs from Step 2 (\widehat{D}_m) are put into the BN model to calculate the relative level of risk. The network structure follows Figure 3, parameters are presented as nodes in the BN. A mechanism to inject risk perspective into the conditional probability tables (CPTs) (see an example in Table 2) is necessary. The BN should be able to reflect unequal importance of different risk parameters according to different risk management strategies or companies' situations. Denote the probability of a node at the state m (Low, Medium, and High) as p_m . p_m can be calculated by adding the weight of all parent nodes (lower-ranked parameters) at the same state w_{pm} (Equation 8). A similar interpolation method was proposed by Alyami et al. (2014), Nguyen and Wang (2018), and developed to be customizable by Wan et al. (2019) and Nguyen et al. (2019). For example, assume $w_L = x$; $w_S = y$; $w_D = z$ with $x + y + z = 1$. The CPT of R when L is Low is shown in Table 2.

$$p_m = \sum w_{pm} \quad (8)$$

TABLE 2 The CPT of Risk level (R) with Likelihood of occurrence (L) is Low

L	Low								
S	Low			Medium			High		
D	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	1	$x + y$	$x + y$	$x + z$	x	x	$x + z$	x	x
Medium	0	z	0	y	$y + z$	y	0	z	0
High	0	0	z	0	0	z	y	y	$y + z$

The probability calculation of the risk level follows the Bayes' theorem. For example, the probability of the child node A in state a , ($a = 1,2,3$), denoted by $p(A_a)$, whose parent nodes are B and C , is calculated as Equation 9 (Yang et al., 2008).

$$p(A_a) = \sum_{b=1}^3 \sum_{c=1}^3 p(A_a|B_b, C_c) p(B_b) p(C_c) \quad (9)$$

By getting the final probability distribution of risk level $p(R_r)$, Risk Index (RI) value of risk i can be computed by using utility values (V) for states as Equation 10 (Alyami et al., 2014). Here, a logarithmic scale is defined as $V_1 = 10^0$; $V_2 = 10^1$; $V_3 = 10^2$ to express the criticality of each state. An HE with a higher RI is considered having a higher risk level.

$$RI_i = \sum_{r=1}^3 p(R_r) V_r; (i = 1, 2, \dots, Y) \quad (10)$$

3.5. Risk mapping

In this study, UI is considered as the key concept in providing insights into the strength of the knowledge base for each CSOR. The response of the system manager after the risk prioritization process is a focus here: Could individual RMP plans be developed immediately, or more data and knowledge should be collected beforehand, or both, in parallel? A interpretative method to support decision-making will be beneficial in this situation. A risk map is developed in a 2-dimensional space of RIs and UIs (Figure 4). By identifying the most enigmatic risks, shipping companies could continuously investigate their operational risk situation and be able to make well-informed RMP decisions.

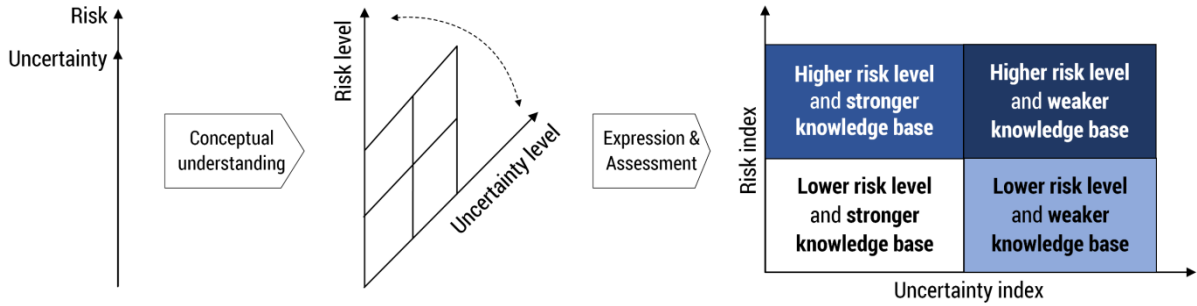


FIGURE 4 Risk mapping with qualitative interpretations of areas

4. VALIDATION PROCESS OF THE PROPOSED CSOR ASSESSMENT MODEL

The validity of a QRA could be divided into two primary aspects: *conceptual and foundational*; and *pragmatic* (Goerlandt et al., 2017). In this study, an empirical case study was conducted, followed by an evaluation of the validity evaluation, and finally the investigation on the sensitivity of both the BN network and the complete model.

4.1. The application of the proposed method – A case study

While risk quantification studies are usually reasoned and supported by scientifically proved theories, the missing of actual applications negatively impacts their reliability (Aven and Zio, 2014, Goerlandt and Montewka, 2015b). A case study was conducted in a container shipping company to examine its functionality, feasibility, and advantages in filling the mentioned gaps in Section 2.5.

4.1.1. Risk identification

Based on the introduced risk concept, risks in container shipping operations could be recognized in the form of potential HEs. Activities in container shipping can be divided into three logistics flows: Information, Physical, and Payment. Focusing on the QRA model, this study utilizes the CSOR list

provided by Chang et al. (2015) and Nguyen and Wang (2018) in Table 3 because of two main features. *Firstly*, they identified typical CSORs from the particular view of shipping company, thus compatible with the targeted user of our proposed QRA model. *Secondly*, these studies described risks in the form of potential HEs with clarifications of causing factors and possible consequences, hence compatible with the risk concept of our model ($R = A, C, U$) and avoid ambiguity of experts in understanding the HE scenarios to be assessed. Risks are coded based on their original flows and categories.

TABLE 3 List of CSORs implemented for the case study

Flow	Category	Potential HEs	Risk code
Information	Information delay	Differences between communication interfaces	ID1
		Lateness in information transmittance from partners	ID2
		Unexpected postponement of compulsory formalities	ID3
	Information incompleteness or inaccuracy	Cyberattacks on the connected IT systems	II1
		Incompleteness and asymmetry of information received	II2
		Incompatibility between implemented IT systems	II3
		Additional information required by partners or customers	II4
		Misdeclaration of cargoes in containers	II5
	IT technical failures	Erroneous operations of IT components (system deficiencies)	IT1
		Human-caused errors on the IT systems	IT2
Physical	Transportation delay	Halts of port due to strike, unrest or war situation	TD1
		Port or terminal congestion	TD2
		Low port/terminal container handling productivity	TD3
		Unforeseen slow steaming or stoppage of vessel	TD4
		Shortage of empty container	TD5
		Rigid operation and management of the fleet	TD6
		Detainment of shipments or ships by authorities	TD7
		Acts of God	TD8
	Loss/damage of goods/assets	Inland transport deficiencies (e.g., accidents, congestion)	TL1
		Maritime accidents (include inland waterway)	TL2
		Accidents in cargo handling operations	TL3
		HEs caused by dangerous goods (e.g., leak, fire, explosion)	TL4
		Failures in maintaining reefer containers' temperature	TL5
		Shipments being stolen or container being tampered with	TL6
		Acts of piracy and terrorism in the maritime legs	TL7

Payment	Payment delay	Payment delays by partners or shippers	PP1
		Unrealized contract or agreement	PP2
	Decrease or total loss of payment	Fluctuation of currency exchange rate	PD1
		Unexpected rise of fuel costs	PD2
		Financial difficulties or bankruptcy of partners or customers	PD3
		Breaks of contract and/or reduce of the booking volume	PD4
		Having partners or customers with low credibility	PD5
		Abandonment of containers at the port of destination	PD6

4.1.2. Risk assessment and result analysis

A case study was conducted on a shipping company in Haiphong, Vietnam. The company has container shipping and logistics integrated services as one of its core businesses. Shipping and consolidation services for customers in the northern industrial zones is also an important segment of the company. Regarding the transport capacity, the company has five feeder-size container ships with the gross capacity of more than 5000 TEU. The fleet is fed by container barges and semi-trailer trucks operated by both the company and its outsourcing partners. Six experts were recruited for the input extraction phase (Table 4). A broad range of experts with diversified background experience is recommended in application of this model, especially with companies that consist of higher number of departments involved to ensure a certain extent of intersubjectivity and the larger knowledge base. The scope of the conducted analysis covers all the processes related to container shipping operations of the company including booking, hauling, consolidation, transshipment, and maritime transport. The calculation tasks were performed on MATLAB and Hugin Expert software.

TABLE 4 Expert's position and experience background in the case study

Department	Position	Position experience and professional background
Operation and Trade: work directly with customers, partners and managing related operations of the fleet	Vice manager	17 years of experience in the shipping industry, 14 years as a mariner with 3 years as captain of 2 vessels
	Sale coordinator	11 years worked in 2 shipping agencies before joining the current company
Financial and Accounting: in charge of financial affairs including tax, payment, portfolio and payroll management	Department manager	21 years of working experience in 3 shipping companies. 5 years in the current position
	Accounting assistant	12 years of experience in different accounting positions of the current company

Technical and Supply: in charge of technical support and supply managing for onshore and offshore operations	Deputy manager	13 years of experience working in the parent company as a mariner, 4 years as a chief engineer and 4 years in the current company as a technical advisor
	IT specialist	11 years in the field with 3 years in the current position. Specialized in maintenances and operations of IT system in the company

Step 1: Input extraction – CSOR survey and data collection process

Since there is no significant difference between the experts in terms of their experience background and no substantial evidence was found in this aspect, the weights of experts were considered as equal for this case study (see Rae and Alexander (2017) for mechanisms that constitute “expertise” in risk assessment). A questionnaire was designed to collect assessments from these experts (see Section 3.1 for the description of the process). The definitions for individual states of risk parameters (Table 5) were improved to be more exclusive and graspable based on the studies of Nguyen and Wang (2018), Wan et al. (2019), and Nguyen et.al, (2019). The thresholds (e.g., the range of financial impact) were discussed and determined deliberatively by the expert panel to suit the perspective of the current company. The estimation of UDoBs in Table 6 is based on the framework of Goerlandt and Montewka (2015a) and agreed upon by the experts before the input extraction process. To ensure relatively consistent reasoning pathways have been used by experts to deduce risk assessments, a second round of questionnaire was conducted after the first round, separated by a short time period (days). The consistency in the assessments of experts through two rounds is monitored individually. Experts whose assessments exhibit >15% disparity were informed to give their final verdicts. The rest of the final database would be filled with arithmetic averages of Round 1 and 2. The rate of inconsistency detected is as follows, Expert 1: 9.34%, Expert 2: 11.74%, Expert 3: 5.18%, Expert 4: 13.38%, Expert 5: 11.49%, Expert 6: 15.53%. These rates are computed based on the total of $33 \text{ risks} \times 6 \text{ parameters} \times (3 \text{ DoBs} + 1 \text{ UDoB}) = 792$ subtasks of assessment.

TABLE 5 Concepts of states for assessment of risk parameters

Parameter	State of assessment		
	Low	Medium	High
L	<p>The HE is considered as unlikely to occur</p> <p>The frequency of events can be enumerated as equal or less than once every six months</p>	<p>The HE is considered as likely to occur but only occasionally</p> <p>The range of frequency is from twice every six months to once every two months</p>	<p>The HE is considered as likely to occur frequently</p> <p>The frequency is twice every two months or more often</p>

F	<p>All the below conditions are met:</p> <p>Minor or insignificant cut considering the revenue from freight</p> <p>Financial damage of less than 12,000 USD/event</p> <p>Negligible negative effect on the solvency or liquidity of the company</p>	<p>One of the below conditions is met:</p> <p>Significant cut considering the revenue from freight</p> <p>Financial damage ranges from 12,000 to 50,000 USD/event</p> <p>Noticeable negative effect on the solvency or liquidity of the company</p>	<p>One of the below conditions is met:</p> <p>Substantial cut or deficit considering the revenue from freight</p> <p>Financial damage of above 50,000 USD/event</p> <p>Substantial negative effect on the solvency or liquidity of the company</p>
I	<p>All the below conditions are met:</p> <p>No noticeable damage on the image of the company</p> <p>No noticeable effects on the relations with partners and customers</p> <p>No observable impacts on the stock price</p>	<p>One of the below conditions is met:</p> <p>Negatively affected image of the company but recoverable in short term</p> <p>Noticeable effects on the relations with partners and customers (e.g., received multiple or major complaints)</p> <p>Observable decreases in stock price</p>	<p>One of the below conditions is met:</p> <p>Considerably affected company's image and credibility</p> <p>Substantial effects on the relations with partners and customers (e.g., discontinuation of cooperation, breaking contracts or agreements)</p> <p>Significant decreases in stock price</p>
O	<p>All the below conditions are met:</p> <p>Negligible disruptions with insignificant negative effects on the operations of the fleet</p> <p>No HR rearrangement or escalation</p>	<p>One of the below conditions is met:</p> <p>Noticeable disruptions with significant negative effects on the operations of the fleet (e.g., skip of ports)</p> <p>Minor rearrangement or escalation of HR</p>	<p>One of the below conditions is met:</p> <p>Substantial disruptions with significant negative effects on the operations of the fleet (e.g., halts of vessel, line, and company business)</p> <p>Major rearrangement or escalation of HR</p>
U	<p>All the below conditions are met:</p> <p>High chance of reliable early warning before the HE</p> <p>Well-prepared by effective and officially implemented detection methods or leading indicators (e.g., protocols, regulations) before operation, leg, or voyage</p>	<p>One of the below conditions is met:</p> <p>Difficult to foresee or detect the HE before its occurrence</p> <p>The detection methods or leading indicators are known but not officially implemented or implemented but not effective</p>	<p>One of the below conditions is met:</p> <p>Impossible or extremely difficult to foresee or detect the HE reliably before its occurrence</p> <p>The detection methods or leading indicators are not known or not implemented</p>

T	All the below conditions are met:	One of the below conditions is met:	One of the below conditions is met:
	High chance of timely detection of the HE after it occurred	Unlikely to be detected right after the HE occurred.	Extremely rare to be detected right after the HE occurred
	The HE is likely to be early detected when impacts might be already developed, but the severity is still minor and considered as easily recoverable	The symptoms of the consequences are usually subtle but can be detected during the casual operation	Very likely to progress in a deliquescent manner, can only be discovered and responded with active investigation
		Likely to be detected and responded when the HE has already caused medium level severity of consequences	Likely to be detected and responded when the HE has already caused high level severity of consequences

TABLE 6 The concept of ranges used for UDoB estimation

UDoB range	Range description
[0.8;1)	Experience are not available. The key factors affecting the states of parameters and the development of scenarios are not yet or poorly understood, causing severe hesitation in giving assessments. Information regarding the HE, affecting factors, and possible scenarios are required before any act of RMP
[0.6;0.8)	Data are only available in sporadic mentions in reports. The key affecting factors and the related phenomena are poorly understood; the development of scenarios is difficult to cover, causing significant hesitation in giving assessments. More insights regarding occurred instances (e.g., industrial peer examples, internal investigations) as well as reasoning of likely scenarios are strongly recommended before RMP
[0.4;0.6)	Data are available but are incomplete, incompatible with the current parameters, or limited to brief reports. The key affecting factors and the related phenomena are somewhat understood, but the development of scenarios are only dubiously conjecturable, causing noticeable hesitation in giving assessments. More insights regarding likely scenarios are recommended before RMP. Risk RMP can be planned and implemented but with high cautions
[0.2;0.4)	Data are available and compatible with the current parameters. The key affecting factors and the related phenomena are reasonably understood. The development of scenarios are

	conjectured with a sense of confidence. Risk RMP can be planned and implemented with degrees of confidence
[0;0.2)	Data are available, compatible with the current parameters, and recorded with details. The key affecting factors and the related phenomena are well understood. The development of scenarios are conjectured with high confidence. Risk RMP can be planned and implemented with confidence

Experts also gave their assessments regarding their perceived importance of the risk parameters. Various weighting methodologies could be used for this purpose (Wang and Nguyen, 2016). In this case study, a simple 1-5 Likert scale in combination with linguistic variables (Very Low, Low, Medium, High, and Very High) was used. It could be interpreted from the results in Table 7 that the panel in this case study leaned toward a more “reactive” risk perspective (Safety-I) where risk likelihood and consequences (L and S) were considered as more significant than risk detectability and responsiveness (D). Financial impact (F) was dominant in different types of consequences while the self-concealing character of risk (U) had higher influence than the lateness of detection (T) on the overall risk level (R).

TABLE 7 Weights of risk parameters

Criteria	L	S	D	F	I	O	U	T
Notation of weights	w_L	w_S	w_D	w_F	w_I	w_O	w_U	w_T
Weights	0.3	0.4375	0.2625	0.4396	0.2637	0.2967	0.5238	0.4762

Step 2: Risk assessments aggregation

After the input extraction process, an ER algorithm was applied to aggregate risk assessments. An example of aggregation for the criteria Financial impact (F) of the HE ID1 is illustrated in Table 8. While the *normalized aggregated DoBs* (\widehat{D}_m) are processed as inputs for the BN model, the *total expert ignorance* ($\sum DU$) will be used in combination with the *average disagreement* (Δ_A) and *total assessment polarization* ($\sum \Delta_P$) in calculating UI in the next step.

TABLE 8 An example of data aggregation and normalization

	DoBs			UDoB
	Low	Medium	High	
Expert 1	0.6	0.2	0.05	0.15
Expert 2	0.5	0.15	0.05	0.3
Expert 3	0.5	0.15	0.1	0.25
Expert 4	0.475	0.225	0	0.3

Expert 5	0.6	0.2	0.1	0.1
Expert 6	0.6	0.2	0.15	0.05
Aggregated	0.5927	0.1700	0.0640	0.1733
Normalized	0.7169	0.1787	0.0668	0.1733

Step 3: Risk and knowledge base assessment

Equations 3-5 were used to calculate the normalized values $\hat{\Delta}_A$, $\hat{\Sigma}\hat{D}U$, and $\hat{\Sigma}\hat{\Delta}_P$. These data are the inputs of Equations 6 and 7 to compute the UI value of each CSOR. Regarding RI, all the CPTs employed in the BN model were built by the normalized weights of risk parameters (Table 7) and applying Equation 8. For example, the CPT of R when L is Low was constructed as in Table 9. \hat{D}_m were put into the Equation 9 (BN model) where the final DoBs could be derived. Finally, the RI value of each risk was calculated using Equation 10. An example of the risk code ID1 is illustrated in Figure 5. $RI_{ID1} = 0.7129 + 0.2057 \times 10 + 0.0814 \times 100 = 10.9099$.

TABLE 9 The CPT of Risk level (R) with Likelihood of occurrence (L) is Low

L	Low								
S	Low			Medium			High		
D	Low	Medium	High	Low	Medium	High	Low	Medium	High
Low	1	0.7375	0.7375	0.5625	0.3	0.3	0.5625	0.3	0.3
Medium	0	0.2625	0	0.4375	0.7	0.4375	0	0.2625	0
High	0	0	0.2625	0	0	0.2625	0.4375	0.4375	0.7

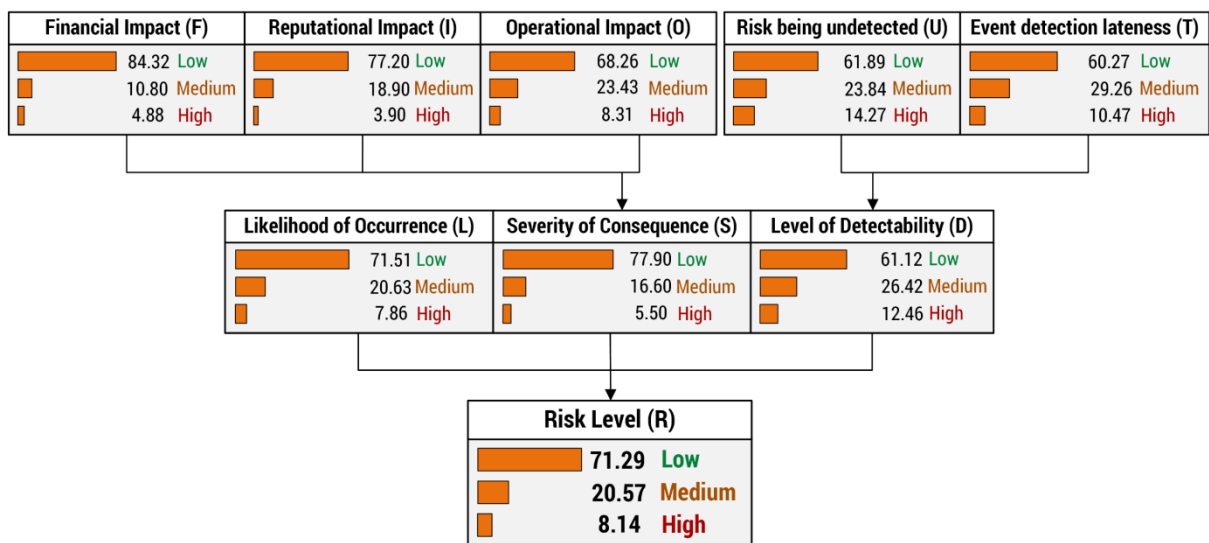


FIGURE 5 The BN model for assessment of the risk ID1

Step 4: Risk prioritization and mapping

After the calculation of all RI values, a prioritized list of identified CSORs was drawn up as in Table 10. Overall, the model proved its separability with no coequal risks in the list. In this case, the dominance of physical risks as the most critical CSORs were well-observed with 60% of the five and 50% of the ten most critical risks originated from this flow. The ranked first three physical are from the *Loss/damage of goods/assets* category including HEs that actually inflicts damages on the cargoes or the ships: *Acts of piracy and terrorism in the maritime legs* (TL7); *Maritime accidents (include inland waterway)* (TL2) and *HEs caused by dangerous goods* (TL4). While these risks were assessed by experts as not likely to occur (very high probability of the “Low” state), their financial impact can be tremendous even when parts of the damages are compensable by insurance and, their effects on the company’s credibility and the continuity of its operations can be substantial. This result also highlights the ability of the proposed QRA model in addressing the risks that have very high consequences but very low likelihood of occurrence, which might be underestimated by the traditional method of ($Risk = Probability \times Consequences$) (Rosqvist, 2010, Berle et al., 2011, Ale et al., 2015).

TABLE 10 Prioritized list of CSORs based on RI values

Risk code	Hazardous event						RI value	Rank
TL7	Acts of piracy and terrorism in the maritime legs						42.0799	1
TL2	Maritime accidents (include inland waterway)						40.9874	2
PD2	Unexpected rise of fuel costs						40.6989	3
II5	Misdeclaration of cargoes in containers						39.7401	4
TL4	HEs caused by dangerous goods (e.g., leak, fire, explosion)						39.3130	5
II2	Incompleteness and asymmetry of information received						35.9821	6
TD2	Port or terminal congestion						34.5352	7
PD6	Abandonment of containers at the port of destination						33.8768	8
TL3	Accidents in cargo handling operations						33.5519	9
PD3	Financial difficulties or bankruptcy of partners or customers						32.4497	10
TD3	Low port/terminal container handling productivity						32.3167	11
TD7	Detainment of shipments or ships by authorities						30.4461	12
PP1	Payment delays by partners or shippers						29.7495	13
IT2	Human-caused errors on the IT systems						29.3162	14
TD6	Rigid operation and management of the fleet						28.5719	15
Risk code	RI value	Rank	Risk code	RI value	Rank	Risk code	RI value	Rank
TL6	27.9767	16	TD4	23.5666	22	TL5	20.1679	28
PD1	26.4430	17	TD1	23.3209	23	ID2	19.6075	29
TL1	25.1808	18	TD8	23.1889	24	PD5	19.5819	30
PP2	24.6042	19	IT1	21.3563	25	PD4	17.9178	31
II4	24.2683	20	ID3	21.0361	26	II3	12.9350	32
II1	23.6281	21	TD5	20.2416	27	ID1	10.9104	33

The most crucial payment and information risks are *Unexpected rise in operational costs* (PD2) and *Shippers hiding cargo information (non-declare)* (II5) (Table 10). While the most critical CSORs in the information flow belong to the “*Information incompleteness or inaccuracy*” category (II5, II2), the contribution of the “*Decrease or total loss of payment*” in the payment flow is observable (PD2, PD6, and PD3). II5 attracted significant concerns because of its “deliquescent” nature of related accidents. It was described as a great risk where dreadful results could happen when the dangerous characteristics of cargo were not fully or deceptively described, hence safety procedures such as safely stuffing, lashing, and segregation cannot be applied. High frequency of recent accidents with chemical substance shipments onboard container vessels or terminals such as *Maersk Honam*, *Yantian Express* and *Tianjin Port* are well-observed (Leander and Lin, 2015, Porter, 2018). In the payment flow, while PD2 is able to cause a serious damage to the profitability of the fleet, adaptive solutions to this risk are still limited. Although *fuel hedging* might appear as a promising contractual tool for stabilizing against fuel price fluctuations, forecasting failures can cause backfires with even larger financial damages. Another strategy is applying bunker surcharge, which is observable in the industry as experiencing protests by major shippers – causing reputational damages.

TABLE 11 Ranked list of CSORs based on UI values

Risk code	Hazardous event						UI	Rank
II5	Misdeclaration of cargoes in containers						78.3637	1
PD2	Unexpected rise of fuel costs						73.7543	2
II2	Incompleteness and asymmetry of information received						73.1525	3
IT1	Erroneous operations of IT components (system deficiencies)						71.6836	4
IT2	Human-caused errors on the IT systems						71.2873	5
TD4	Unforeseen slow steaming or stoppage of vessel						70.6216	6
TD6	Rigid operation and management of the fleet						65.8434	7
PD1	Fluctuation of currency exchange rate						62.3990	8
TD5	Shortage of empty container						61.2888	9
TD3	Low port/terminal container handling productivity						55.2797	10
PP1	Payment delays by partners or shippers						54.9113	11
II1	Cyberattacks on the connected IT systems						54.2196	12
TL6	Shipments being stolen or container being tampered with						54.1807	13
TL3	Accidents in cargo handling operations						53.6413	14
TD8	Acts of God						53.0439	15
Risk code	RI value	Rank	Risk code	RI value	Rank	Risk code	RI value	Rank
PD3	49.1025	16	ID2	42.8052	22	PD4	35.1870	28
PP2	46.9143	17	TD1	42.0003	23	TD7	35.0438	29
TD2	46.8868	18	TL7	39.5214	24	TL2	34.4575	30
II3	46.5422	19	PD6	39.2905	25	ID3	34.3154	31
II4	46.4312	20	PD5	36.7438	26	TL4	28.6791	32

TL1	43.0290	21	TL5	36.2375	27	ID1	24.2138	33
-----	---------	----	-----	---------	----	-----	---------	----

The results of uncertainty assessment are presented in Table 11, which shows a different prioritization result in comparison with that of risk assessment. The usefulness of recognizing and actively assessing the uncertainty dimension is observed through three aspects. *Firstly*, deeper investigation is helpful in tracking down and filling in knowledge weaknesses in a prioritized order. The dedicated UQM module was able to pinpoint the improvable “blind spots”. For example, CSORs from the information flow occupied 4/5 highest positions (II5, II2, IT1, and IT2) while the risks of transportation delay (TD4, TD6, TD5, TD3) and payment damage (PD1, and PD2) were assessed with high uncertainty. *Secondly*, the UQM model was able to provide deeper insights into the symptoms and causes of the perceived weak knowledge base. For instance, the higher uncertainty in various *transportation delay* risks suggests a possible systematic shortage of historical data related to experienced frequency and time delayed, which is the main factors influencing their likelihood and consequences. *Thirdly*, the heatmap in Figure 6 indicates a counter-bias mechanism where different indicators were utilized in cross-checking symptoms of inadequate knowledge base. PD2 and II2 were decisively agreed among the experts as having high level of uncertainty (low Δ_A , $\Sigma\Delta_P$, and high ΣDU). Meanwhile, the RI values of TD1 and TD4 was assessed with confidence (low ΣDU) but the expert judgements were significantly dissimilar (high Δ_A), exhibited a degree of overconfidence in the QRA result.

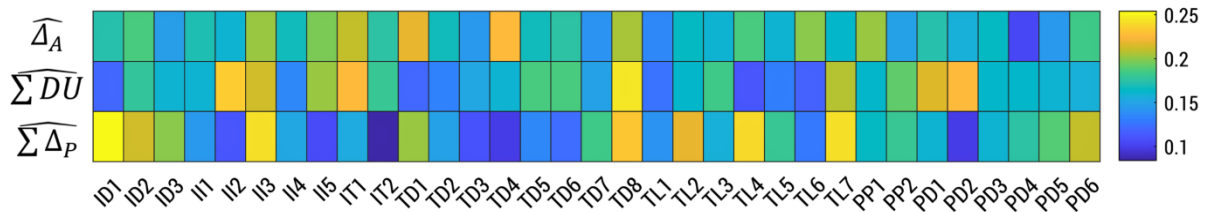


FIGURE 6 The heatmap of normalized uncertainty indicators of the case study

The analysis of the UQM in the context of CSOR is meaningful in two ways towards QRA models that involve subjective probabilities from experts. *Firstly*, the knowledge base is a critical aspect in assessing risk that is worth attention of risk analysts. The case study showed a diversified situation of knowledge base’s quality that might be critical if risks were treated out of their proportion (under- or overestimation) due to knowledge base deficiencies. *Secondly*, relying solely on experts to assess the knowledge base (UDoBs) can overlook other weaknesses of it (e.g., ambiguity, discrepancies in reasoning). This limitation is explainable by (1) the fact that UDoBs are only estimated based on individual knowledge bases, and (2) the subjectivity of expert in self-assessment, even with the support of a clear framework (i.e., description of UDoB scale in Table 6). This phenomenon validated the usefulness of the UQM that can provide a more reliable evaluation of the expert panel’s overall knowledge base.

The values of RI and UI are visualized in a 2-dimensional space to provide a graphic relative view of both risk level and the strength of knowledge base (Figure 7), which has three main features. *Firstly*,

there are HEs that have relatively low RI and UI values such as ID1, PD4, PD5, and TL5 in Zone 3. These CSORs have low priorities and relatively low uncertainty, thus could be considered as temporarily ignorable and should be put into a watch list since the resources of the company should be better distributed to other risks. *Secondly*, CSORs with high-RI and low-UI such as TL4, TL2, TD7, TL7, and PD6 were assessed with significant confidence. Their RMP strategies can be developed immediately. *Finally*, the chance of type-II error with low-RI risks such as II1, TD5, IT1, and TD4 in Zone 4 or type-I error with high-RI risks such as IT2, II2, II5 and PD2 in Zone 2 could be lowered by strengthening the knowledge base through further investigation into the influencing factors and potential scenarios. There are different strategies for that purpose such as encouraging the interaction of experts in the panel (deliberative assessment), using third-party shipping financial consultants, increasing the number of experts in the panel, and obtaining more critical data about individual risks from the flow. As being demonstrated, the proposed risk mapping method is helpful in suggesting the needed immediate actions for CSORs in an orderly manner.

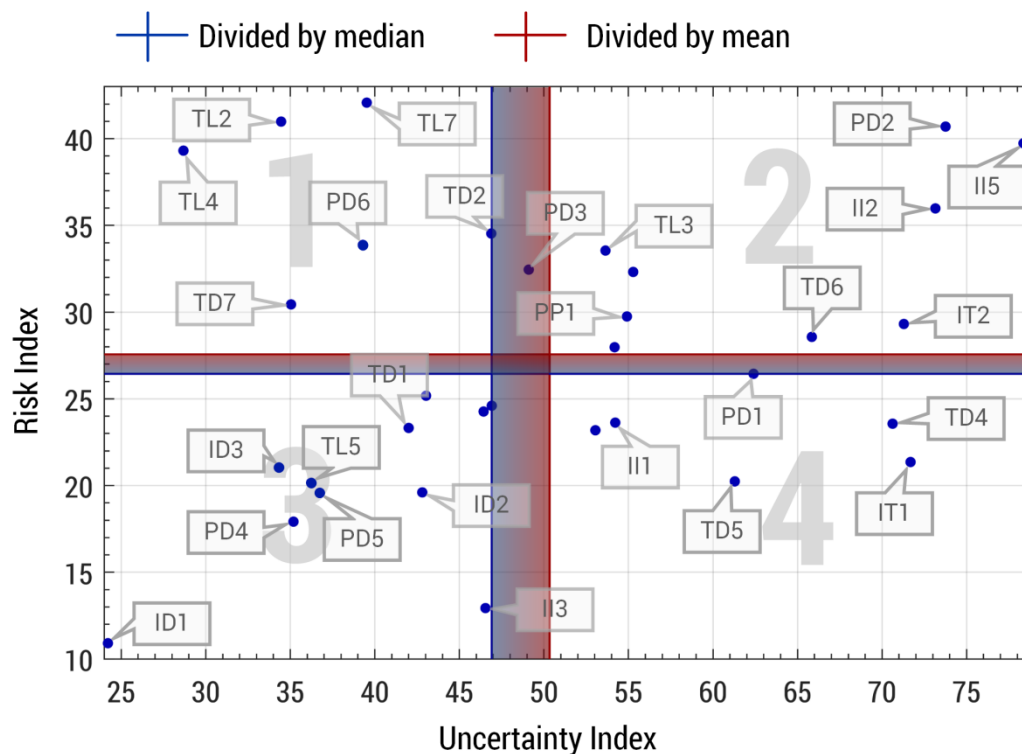


FIGURE 7 Risk map generated from the case study

4.2. The case study's evaluation of validity

CSORs continuously change under multiple dynamic factors of the industry (e.g., human, natural, policies, involved parties) and causing highly unique HEs, even under the same risk (i.e., low repeatability). The confrontation of risk assessments with the realistic data collected afterward, therefore, is difficult to be considered as yielding the QRA's validity. This problem with the validation QRA model and the criticisms toward the accuracy claim have been argued by Aven and Heide (2009)

and Goerlandt et al. (2017) extensively. Based on the conducted case study, the QRA model in this study was checked for multiple validity criteria mentioned in the literature (Table 12). The derived results are displayed in Table 13 with all the aspects were attained in the case study, proving that the proposed QRA model is capable of assessing CSORs in a descriptive and reliable manner.

TABLE 12 Recommended factors for validation of the proposed risk assessment method

Category	Aspect	Explanation
Conceptual and foundational	Cognitive assessment transformability	The degree to which the assigned subjective probabilities adequately describe the assessor's uncertainties of the unknown quantities considered (Aven and Heide, 2009)
	Uncertainty treatment completeness	The degree to which the uncertainty assessments are complete (Aven and Heide, 2009)
	Quantities addressing rightness	The degree to which the analysis addresses the right quantities (Aven and Heide, 2009)
Pragmatic	Methodologies	The clear explanation and adequate compatibility of the proposed method (Okrent et al., 1982, Rae et al., 2014)
	Assessor panel	The expertise and constitution of the assessor panel (Rae and Alexander, 2017)
	Supporting database	The compatibility and transparency of the supporting database (Okrent et al., 1982, Rae et al., 2014)
	Calculation process	The correctness and transparency of the computing process
	Results and interpretation	The usability and accessibility of the users to the results in the decision-making process (Okrent et al., 1982, Rae et al., 2014)

TABLE 13 Validity evaluation of the conducted case study

Aspect	Evidence
Cognitive assessment transformability	Experts are able to provide (1) <i>risk assessments</i> that, by design of the model, cover the whole range of possible values; (2) <i>uncertainty assessments</i> , through the concept of UDoB in risk assessments.
Uncertainty treatment completeness	The uncertainty awareness of the proposed method is explained in Section 2.2 and the UQM is presented in Section 3.3. The UQM was illustrated in the case study with the handling of both U_O and U_E
Quantities addressing rightness	The experts were able to produce subjective assessments with high consistency rate (~80%) with the introduced processes and the model configurations in the conducted case study

Methodologies	The proposed methodologies are combined based on both risk theoretical understandings and their applicability to the container shipping industry. The model's mathematical reliability is validated by a sensitivity test
Assessor panel	Experts with experience and managing positions in the company was recruited for the case study to ensure both <i>private information</i> and <i>domain knowledge</i> mechanisms of expertise (see Rae and Alexander (2017) for their explanations)
Supporting database	Insights about the supporting database were gained through the proposed UQM featured three indicators. The case study proved the UQM's usefulness
Calculation process	Calculation processes were carried out automatically on MATLAB and the .NET APIs of Hugin Expert software to ensure the computation accuracy ¹
Results and interpretation	Both risk level and the strength of the knowledge base were communicated in the final risk map. The case study illustrated the usefulness of risk map in assisting the decision-making process with informative visualization

4.3. Sensitivity test

A sensitivity test is designed to generalize the effectiveness of the proposed QRA model to the whole set of possible inputs. By adjusting input values followed by observation of the corresponding RI and UI values, insights into the mathematical reliability of the model can be gained. Four axioms below are employed. Three first three axioms are based on the study of Yang et al. (2008) while the fourth axiom is added by the authors to take the impacts of the agreement among experts into account.

Axiom 1. Any adjustment of the input certainly results in a relative variation in the output of the model, including RI and UI values. This axiom ensures the power of the model in distinguishing multiple risks.

Axiom 2. The total influence magnitudes of multiple input adjustments that have the same effects (positive or negative) on the output are always greater than any one of its subsets. This axiom ensures the significance of all inputs as well as the rightness in the modelling process.

Axiom 3. The tendency and degree of variations in the output with any adjustment of the input align with the expected influencing mechanisms. This axiom ensures the rightness in the quantification of the risk perception.

Axiom 4. The fluctuations of output caused by the same input adjustment are in accordance with degrees of agreement between experts. This axiom ensures the robustness of the model in dealing with unilateral or minor fluctuations of the input.

¹ MATLAB® written functions can be provided upon request

4.3.1. BN sensitivity test

Probability was changed with steps of 1% from an “absolute low” assessment (100, 0, 0), to “absolute medium” (0, 100, 0) and finally to “absolute high” (0, 0, 100). The derived RI values are monitored along the process and shown in Figure 8.

Figure 8 demonstrates that the BN model satisfied *the first axiom*. The magnitudes of affecting effects of the parameters on the final RI are indicated by the slope of the lines. All RI values increased considerably sharper after the 100th case since the attached utility value (V_r) was raised from 10 to 10^2 (Equation 10). The RI values grew linearly with the adjustments of the input data with all parameters. “Hard evidence” was arbitrarily set on S and D to check their influential magnitudes on R . The results in Figure 8 also follow the relative weights of parameters (Table 7), thus satisfies *Axiom 2*. *Axiom 3* is satisfied based on the fact that the “all input nodes” scenario has the highest impact; and both S and D have larger impacts than their secondary parameters. These experiment results also validate the CPT building method presented in Equation 8. Since the BN model does not consider multiple experts, it was exempted from Axiom 4.

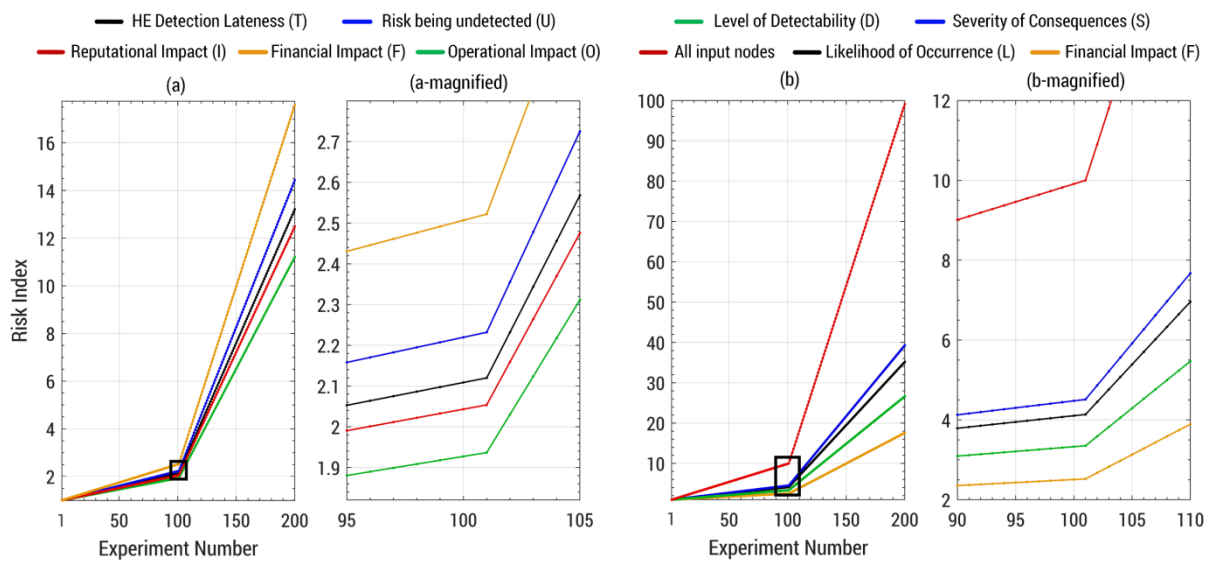


FIGURE 8 The RI values recorded in experimental scenarios with different parameters

4.3.2. Complete method sensitivity test

Another series of experiments were conducted to test the sensitivity of the whole QRA model. The DoB was moved from the Low to Medium based on a baseline distribution of (30, 30, 30, 10), reached (0, 60, 30, 10) and then to High state by 1% per step to the final distribution of (0, 0, 90, 10). The experiment was repeated on all the parameters of the proposed QRA model in the order of ascending importance as shown Figure 8 ($O < I < T < U < F < D < L < S < R$). Three dummy experts were created to investigate the response of the model with extents of agreement in an expert panel. With the exception

of the targets of adjustment, other inputs remained at baseline throughout the experiment. The results of $(91 \times 9) \times 3 = 2457$ cases are illustrated in Figure 9.

Firstly, as the input changed, the experiment results exhibited different results of the UI and RI combination. The indicators of uncertainty (Δ_A , $\sum DU$, and $\sum \Delta_P$) also responded to input adjustments. Hence, the QRA model satisfies *Axiom 1*. Secondly, the influence magnitudes of input adjustments satisfy *Axiom 2*. While the changes of all recorded values in different risk criteria showed similar types of variation, the primary parameters displayed a more impactful role than its secondary ones ($R > L, S$ and $D; S > F, I$, and $O; D > U$ and T). Thirdly, the model behaved as expected regarding the tendency of variations resulted from changes of different input parameters. The degree of fluctuation also aligned with the predefined order based on weighted importance ($O < I < T < U < F < D < L < S < R$). This quality could be observed across all scenarios, thus proves the satisfaction of the proposed QRA model with *Axiom 3*. Finally, as the agreement among experts in the panel increased, the impacts of input adjustment varied accordingly. The growth of RI and decline of UI is obvious in Figure 9. It is noteworthy that while the $\sum \Delta_P$ value raised with the deviation of inputs from the baseline and lowers UI, Δ_A did not. The same Δ_A values were derived in 1/3 and 2/3 experts scenarios since the quantified disagreement among experts was kept unchanged. In the 3/3 experts scenario, Δ_A was reasonably reduced to 0. The reflection of higher confidence by the ER algorithm was also observed as the aggregated value of UDoB was lower than the inputs. However, that effect on $\sum DU$ could only be observable in the case of total agreement (3/3) across the expert panel. Therefore, the proposed QRA model satisfies *Axiom 4* and ultimately, the mechanistic sensitivity test.

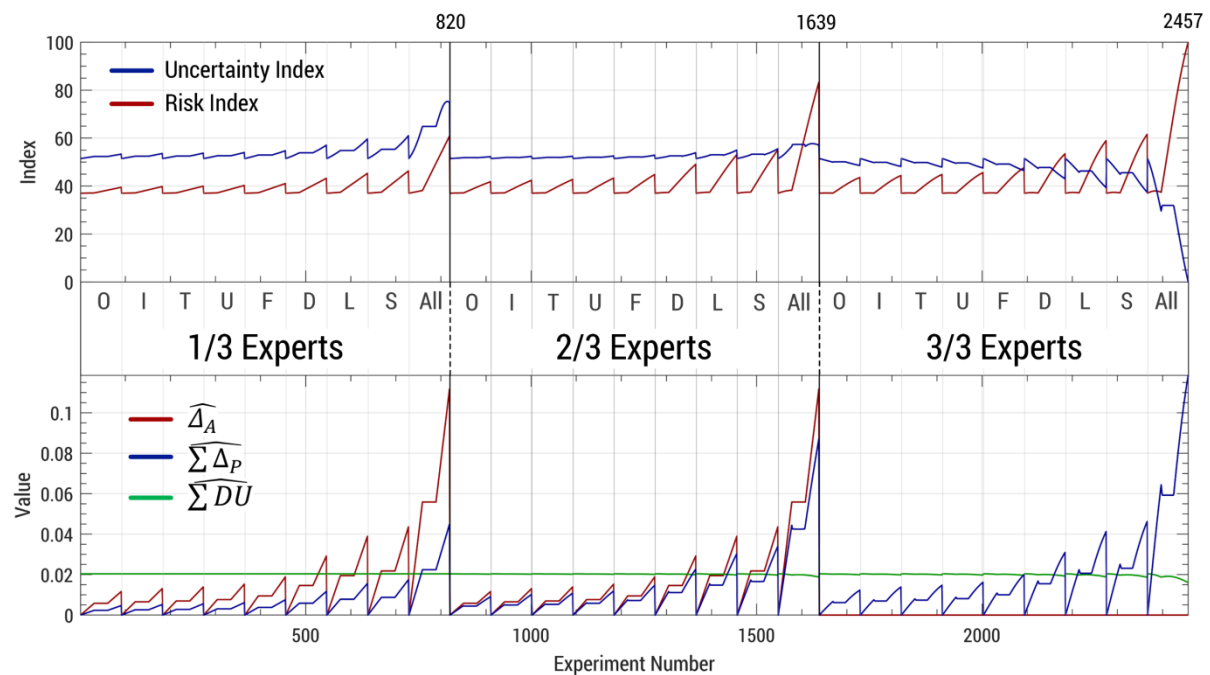


FIGURE 9 Recorded results of indexes and uncertainty indicators in the experiment

5. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

The high volatility of operating conditions and low data availability environment of CSOR make the risk assessments of experts vulnerable to both U_O and U_E . The epistemic nature of uncertainties points toward the strength of knowledge base, which is not yet paid an adequate attention in current managerial QRA models. This study puts forward a solution of quantifying uncertainty through both U_O and U_E in the process of CSOR assessment. The proposed UQM captures the uncertainty situation and facilitates a progressively better QRA through continuous improvement of the knowledge base. A set of three validation processes was applied on the QRA model including sensitivity analysis, validity evaluation, and pragmatic applicability.

The results of this study have three main contributions. *Firstly*, this study proves that the recognition and awareness of uncertainty, especially in the situation of low data availability, is critical with the reliability and significance of QRAs. The realization of uncertainty through UI calculation and risk mapping opens another dimension of RMP strategies with the improvement of the knowledge base. *Secondly*, despite being widely applied and well-established, this study suggests there are still rooms for improvement of risk analysis tools and their art of application with updated theoretical risk knowledge. The proposed QRA model attempts to comprehensively capture risk with uncertainty by the reasonable combination of multiple apparatuses. The reliability of the model was justified and validated by a comprehensive theoretical and pragmatic validation process. *Finally*, the case study found the limitations of experts in uncertainty self-awareness and the instances of CSORs that assessed with weak knowledge base. These findings prove the usefulness of the proposed QRA model and the UQM and effectively validate the motivation of this study.

This paper also contributes to the family of QRA models that feature risk prioritization. *Firstly*, a two-level parameter set is reasoned for assessing risks based on a strong connection with the context of CSOR. Expanding to a secondary level with specified aspects of consequence and detectability allows experts to assess risk with a higher level of details and better compatibility with the collected evidence. *Secondly*, the application of ER enables the subjective expression of the strength of knowledge through UDoBs, which is the basis to separate of U_E from U_O and, therefore, supports the complete handling of uncertainty. *Thirdly*, a dedicated UQM is integrated into the QRA model. The awareness of uncertainty will contribute to continuously improving the knowledge base for risk assessment in a prioritized manner. *Finally*, through the visualization of risk map, counter-risk strategies can be planned considering the level of individual risks as well as its current basis of knowledge.

Three potential research directions are identified based on this study. *The first direction* is the utilization and cooperation of factual evidence and the deduction power of the expert panel need deeper and more rigorous investigations. One of the main focuses should be the balance and controllable interaction among experts. It is intuitive and rational to encourage a deliberative and communicative environment

for exchanging risk-rated knowledge. While allowing the beneficial effects of pooling knowledge, brainstorming, or cross-checking, group-based thinking and assessment are also vulnerable to various short comings such as social manipulation and intimidation, and contamination of misleading data. In this respect, the case study of Nguyen et al. (2019) proposed a QRA model in which a Risk Communication Platform (RCP) based on Delphi was used to establish a balance between mathematical and social aggregation. However, the model of Nguyen et al. (2019) has its own limitations that were addressed by the QRA model in this paper. *Firstly*, evaluating the knowledge base more objectively through the UQM with three different indicators revealed the possible subjective bias of the fuzzy self-assessment system. *Secondly*, the current QRA model tolerate a larger and flexible number of assessors. Delphi inherently requires a significant effort of experts in order to stimulate the convergence of the assessments through multiple iterations, which will exponentially increase with the growth of the panel. *The second direction is* to investigate the optimal number of experts in the panel is not yet specified. It is expected that more experts could either contribute useful expertise or more outlier, or “noise” into the analysis. There might be an equilibrium of this trade-off that would be useful for practical implementations. *Regarding the third direction*, identified HEs can be found in practical situations as having causal relationships, signaling for the existence of “key” events that could affect the risk picture of the whole system. For example, in the case of Maersk Honam fire or Tianjin port explosion, the original cause of the HE stemmed from the hiding or missing of critical cargo information. A QRA approach taking these connections into consideration will be beneficial toward the risk RMP processes.

The applicability of the QRA model can go beyond its CSOR context. A range of applications is enabled by the model’s customizability potential that enables its adoptability in other dynamic risk environments. For example, it might be applicable to quantify operational risks of other parties in the container transport network such as logistic providers, port and terminal operators, shippers, and consignees. The model can also be applied to other sectors of transport such as dry bulk and liquid. A computer-based program can be developed based on this model to assist system managers in continuously monitoring the risk and uncertainty situation at hand, enables the ability of risk-informed decision making. Several elements and configurations of the model should be examined for compatibility before such applications. *Firstly*, the set of risk parameters and their weights should be based on the perspective of the risk bearer. For example, an organization might not value financial objectives as important as others (e.g., transport service users that does not operate for the financial benefits, or governmental parties such as customs and port authority). Definitions of different states of parameters, therefore, also need to be revised. *Secondly*, different parties in the transport network might not share a same risk list. A potential HE of a party or a transport sector might not be considered as a risk by another. Hence, risk identification has to be conducted for each case of application.

REFERENCES:

- AKHTAR, M. J. & UTNE, I. B. 2014. Human fatigue's effect on the risk of maritime groundings - A Bayesian Network modeling approach. *Safety Science*, 62, 427-440.
- ALE, B., BURNAP, P. & SLATER, D. 2015. On the origin of PCDS – (Probability consequence diagrams). *Safety Science*, 72, 229-239.
- ALYAMI, H., LEE, P. T. W., YANG, Z. L., RIAHI, R., BONSALE, S. & WANG, J. 2014. An advanced risk analysis approach for container port safety evaluation. *Maritime Policy & Management*, 41, 634-650.
- APELAND, S., AVEN, T. & NILSEN, T. 2002. Quantifying uncertainty under a predictive, epistemic approach to risk analysis. *Reliability Engineering & System Safety*, 75, 93-102.
- ASKELAND, T., FLAGE, R. & AVEN, T. 2017. Moving beyond probabilities – Strength of knowledge characterisations applied to security. *Reliability Engineering & System Safety*, 159, 196-205.
- AVEN, T. 2012. The risk concept—historical and recent development trends. *Reliability Engineering & System Safety*, 99, 33-44.
- AVEN, T. 2013. Practical implications of the new risk perspectives. *Reliability Engineering & System Safety*, 115, 136-145.
- AVEN, T. & HEIDE, B. 2009. Reliability and validity of risk analysis. *Reliability Engineering & System Safety*, 94, 1862-1868.
- AVEN, T. & KROHN, B. S. 2014. A new perspective on how to understand, assess and manage risk and the unforeseen. *Reliability Engineering & System Safety*, 121, 1-10.
- AVEN, T. & RENN, O. 2010. *Risk Management and Governance: Concepts, Guidelines and Applications*, Springer Berlin Heidelberg.
- AVEN, T. & ZIO, E. 2011. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering & System Safety*, 96, 64-74.
- AVEN, T. & ZIO, E. 2014. Foundational issues in risk assessment and risk management. *Risk Analysis*, 34, 1164-72.
- BAKSHI, N., FLYNN, S. E. & GANS, N. 2011. Estimating the Operational Impact of Container Inspections at International Ports. *Management Science*, 57, 1-20.
- BERLE, O., ASBJORNSLETT, B. E. & RICE, J. B. 2011. Formal Vulnerability Assessment of a maritime transportation system. *Reliability Engineering & System Safety*, 96, 696-705.
- CHANG, C. H., XU, J. J. & SONG, D. P. 2014. An analysis of safety and security risks in container shipping operations: A case study of Taiwan. *Safety Science*, 63, 168-178.
- CHANG, C. H., XU, J. J. & SONG, D. P. 2015. Risk analysis for container shipping: from a logistics perspective. *International Journal of Logistics Management*, 26, 147-171.
- COX, L. A., JR. 2008. What's wrong with risk matrices? *Risk Analysis*, 28, 497-512.
- DUIJM, N. J. 2015. Recommendations on the use and design of risk matrices. *Safety Science*, 76, 21-31.

- ELEYE-DATUBO, A. G., WALL, A. & WANG, J. 2008. Marine and offshore safety assessment by incorporative risk modeling in a fuzzy-Bayesian network of an induced mass assignment paradigm. *Risk Analysis*, 28, 95-112.
- FILINA, L. & FILIN, S. 2008. An analysis of influence of lack of the electricity supply to reefer containers serviced at sea ports on storing conditions of cargoes contained in them. *Polish Maritime Research*, 15, 96-102.
- GOERLANDT, F., KHAKZAD, N. & RENIERS, G. 2017. Validity and validation of safety-related quantitative risk analysis: A review. *Safety Science*, 99, 127-139.
- GOERLANDT, F. & MONTEWKA, J. 2015a. A framework for risk analysis of maritime transportation systems: A case study for oil spill from tankers in a ship–ship collision. *Safety Science*, 76, 42-66.
- GOERLANDT, F. & MONTEWKA, J. 2015b. Maritime transportation risk analysis: Review and analysis in light of some foundational issues. *Reliability Engineering & System Safety*, 138, 115-134.
- GOERLANDT, F. & RENIERS, G. 2016. On the assessment of uncertainty in risk diagrams. *Safety Science*, 84, 67-77.
- GREENBERG, A. 2018. *The Untold Story of NotPetya, the Most Devastating Cyberattack in History* [Online]. USA: WIRED. Available: <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/> [Accessed 29/11 2018].
- HWANG, C.-L. & YOON, K. 1981. *Multiple Attribute Decision Making*, Berlin Heidelberg, Springer-Verlag.
- KRIEG, M. L. 2001. A Tutorial on Bayesian Belief Networks. South Australia, Australia: DSTO Electronics and Surveillance Research Laboratory.
- LAM, J. S. L. & LASSA, J. A. 2016. Risk assessment framework for exposure of cargo and ports to natural hazards and climate extremes. *Maritime Policy & Management*, 44, 1-15.
- LEANDER, T. & LIN, M. T. 2015. *Explosions near Tianjin port leave more than 44 dead and disrupt shipping* [Online]. UK: Lloyd's List: Maritime Intelligence. [Accessed 03/11/2018 2018].
- LEE, C. Y. & SONG, D. P. 2017. Ocean container transport in global supply chains: Overview and research opportunities. *Transportation Research Part B: Methodological*, 95, 442-474.
- LEVIN, R. 2005. *Uncertainty in risk assessment – contents and modes of communication*. Doctor of Philosophy Licentiate, Royal Institute of Technology.
- LIU, C.-L., SHANG, K.-C., LIRN, T.-C., LAI, K.-H. & LUN, Y. H. V. 2018. Supply chain resilience, firm performance, and management policies in the liner shipping industry. *Transportation Research Part A: Policy and Practice*, 110, 202-219.
- NGUYEN, S. & WANG, H. 2018. Prioritizing operational risks in container shipping systems by using cognitive assessment technique. *Maritime Business Review*, 3, 185-206.
- NGUYEN, S., CHEN, P. S., DU, Y., SHI, W. 2019. A quantitative risk analysis model with integrated deliberative Delphi platform for container shipping operational risks. *Transportation Research Part E: Logistics and Transportation Review*. (Accepted on August 2019)
- OKRENT, D., APOSTOLAKIS, G., WHITLEY, R. & GARRICK, B. J. 1982. PRA quality and use. Los Angeles (USA): California University

- OTTOMANELLI, M. & WONG, C. K. 2011. Modelling uncertainty in traffic and transportation systems. *Transportmetrica*, 7, 1-3.
- PORTER, J. 2018. *Talks held to stamp out cargo declaration abuses* [Online]. UK: Lloyd's List: Maritime Intelligence. Available: <https://lloydslist.maritimeintelligence.informa.com/LL1124534/Talks-held-to-stamp-out-cargo-declaration-abuses> [Accessed 03/11/2018 2018].
- RAE, A. & ALEXANDER, R. 2017. Forecasts or fortune-telling: When are expert judgements of safety risk valid? *Safety Science*, 99, 156-165.
- RAE, A., ALEXANDER, R. & MCDERMID, J. 2014. Fixing the cracks in the crystal ball: A maturity model for quantitative risk assessment. *Reliability Engineering & System Safety*, 125, 67-81.
- ROSQVIST, T. 2010. On the validation of risk analysis—A commentary. *Reliability Engineering & System Safety*, 95, 1261-1265.
- SHAFFER, G. 1976. *A Mathematical Theory of Evidence*, New Jersey, USA, Princeton University Press.
- STOPFORD, M. 2009. *Maritime Economics: The third edition*, New York, USA, Routledge, Taylor and Francis.
- TSENG, W. J., DING, J. F. & LI, M. H. 2013. Risk management of cargo damage in export operations of ocean freight forwarders in Taiwan. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 229, 232-247.
- WAN, C., YAN, X., ZHANG, D., QU, Z. & YANG, Z. 2019. An advanced fuzzy Bayesian-based FMEA approach for assessing maritime supply chain risks. *Transportation Research Part E: Logistics and Transportation Review*, 125, 222-240.
- WANG, H. & NGUYEN, S. 2016. Prioritizing mechanism of low carbon shipping measures using a combination of FQFD and FTOPSIS. *Maritime Policy & Management*, 44, 187-207.
- WINKLER, R. L. 1996. Uncertainty in probabilistic risk assessment. *Reliability Engineering & System Safety*, 54, 127-132.
- YANG, J. B. & XU, D. L. 2002. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans*, 32, 289-304.
- YANG, Y. C. 2011. Risk management of Taiwan's maritime supply chain security. *Safety Science*, 49, 382-393.
- YANG, Z. L., BONSALE, S. & WANG, J. 2008. Fuzzy rule-based Bayesian reasoning approach for prioritization of failures in FMEA. *Ieee Transactions on Reliability*, 57, 517-528.
- YUEN, K. F., WANG, X., WONG, Y. D. & ZHOU, Q. 2018. The effect of sustainable shipping practices on shippers' loyalty: The mediating role of perceived value, trust and transaction cost. *Transportation Research Part E: Logistics and Transportation Review*, 116, 123-135.
- ZIO, E. 2007. *An Introduction to the Basics of Reliability and Risk Analysis*, Singapore, World Scientific Publishing.