Weight of Evidence Approach to Maritime Accident Risk Assessment Based on Bayesian Network Classifier

Ana Kuzmanić Skelin[a], Lea Vojković[b], Dani Mohović[c], Damir Zec[c]

Probabilistic maritime accident models based on Bayesian Networks are typically built upon the data available in accident records and the data obtained from human experts knowledge on accident. The drawback of such models is that they do not take explicitly into the account the knowledge on non-accidents as would be required in the probabilistic modelling of rare events. Consequently, these models have difficulties with delivering interpretation of influence of risk factors and providing sufficient confidence in the risk assessment scores. In this work, modelling and risk score interpretation, as two aspects of the probabilistic approach to complex maritime system risk assessment, are addressed. First, the maritime accident modelling is posed as a classification problem and the Bayesian network classifier that discriminates between accident and non-accident is developed which assesses state spaces of influence factors as the input features of the classifier. Maritime accident risk are identified as adversely influencing factors that contribute to the accident. Next, the weight of evidence approach to reasoning with Bayesian network classifier is developed for an objective quantitative estimation of the strength of factor influence, and a weighted strength of evidence is introduced. Qualitative interpretation of strength of evidence for individual accident influencing factor, inspired by Bayes factor, is defined. The efficiency of the developed approach is demonstrated within the context of collision of small passenger vessels and the results of collision risk assessments are given for the environmental settings typical in Croatian nautical tourism. According to the results obtained, recommendations for navigation safety during high density traffic have been distilled.

KEY WORDS
~ Maritime collision model
~ Probabilistic modelling
~ Bayesian Network classifier
~ Weight of evidence
~ Bayes factor
~ Probabilistic reasoning

a. University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia
e-mail: akuzmani@fesb.hr
b. University of Split, Faculty of Maritime Studies, Split, Croatia
e-mail: lvojkovic@pfst.hr
c. University of Rijeka, Faculty of Maritime Studies, Rijeka, Croatia
e-mail: dani.mohovic@pfri.uniri.hr

doi:10.7225/toms.v10.n02.w07

Received on: 06.05.2021 / Revised on: 12.07.2021/ Accepted on: 24.08.2021 / Published online: 26.08.2021

1. INTRODUCTION

The expanding character of Croatian nautical tourism has led to a tremendous increase in maritime traffic density, which, in turn, has raised concerns regarding the navigational safety in the Croatian part of the Adriatic Sea basin. The authorities of the Republic of Croatia have adopted a great number of laws, regulations, and orders in the field of navigational safety; however, further work is needed to assess the safety and minimize the risk of accidents due to novel conditions, (Ministry of Sea & Tourism, 2008). Some important directions towards safety assessment include the development of computational risk evaluation approaches and risk factor identification models in accordance with maritime safety analyses and the assessment

process formalized by International Maritime Organization (IMO) into Formal Safety Assessment (FSA) guidelines. FSA is organized into five sets of tasks: 1) hazard identification (HAZID), 2) risk assessment, 3) identification of risk control options, 4) cost/benefit assessment, and 5) recommendations for risk control, (Kontovas & Psaraftis, 2009). Different aspects of human expertise, the maritime domain knowledge and modelling of complex events related to maritime accidents should be unified to support safety risk assessment of FSA, particularly the HAZID, thus contributing towards the ultimate goal of facilitating activities of stakeholders involved in maritime traffic regulation and management. Human knowledge and experience, as well as expert judgments, are the most important sources of information for the risk assessment inline to definitions of FSA.

The mapping of the domain knowledge domain about a risk assessment have become formalized through a volume of diverse scientific computational approaches, such as machine learning and uncertainty analysis. A number of approaches to qualitative and quantitative knowledge modelling have been investigated in maritime domain, but also other environmental modelling and safety assessments domains, as presented in recent literature surveys, (Huang, et al., 2020), (Chen, et al., 2019), (Lim, et al., 2018). Particularly applicable to maritime domain are Bayesian Networks (BNs), (Pearl, 1988), (Pearl, 2000), which have been recognized as an efficient mathematical tool for modelling maritime accidents. Namely, Bayesian Networks (BNs), (Pearl, 1988), (Pearl, 2000), are directed graphical models which provide framework for accident modelling and analysis by supporting the representation of dependencies and interactions of random variables involved in the probabilistic socio-technical system, where random variables are interpreted as causal factors involved in the maritime accident. BNs have been widely applied over the past years to assess a variety of accident types and scenarios, (Hänninen & Kujala, 2012), (Zhang, et al., 2013), (Zhang, et al., 2018), (Baksh, et al., 2018). From an in-depth review of the literature on maritime accident risk models based on Bayesian Network given in (Zhang & Thai, 2016) and (Chen, et al., 2019), it can be observed that the methodological framework required for qualitative and quantitative model development is well defined.

However, challenges are still encountered in the modelling stage and at the inference level, i.e. in the stage of model deployment. Several challenges can be highlighted. Maritime accidents are rare events for which real-world data provides incomplete and insufficient statistical information, leaving a burden of parameter initialization to maritime experts and data engineers involved in the model development. Next, an influence factor identification metric is not clearly defined with respect to quantitative data, thus leading to poor interpretation of model responses. In order to obtain an assessment of factors influencing the behaviour of complex maritime system, the Bayesian network, as an expert system framework founded in data mining and machine learning, should deliver interpretable quantitative and qualitative scores in an inference task.

Specifically, in a BN inference task, the aim is to come up with an assessment of identified hazard factors by their influence. Commonly used quantitative measures of influence are based on sensitivity analysis or one-at-a-time (OAT) analysis for each individual hazard factor, (Hänninen & Kujala, 2012), (Sotiralis, et al., 2016), where influences are calculated as the difference of hard evidence, i.e. as a difference of state values. The output of these analysis reveals the range of change of the target variable of a model - the larger the output change, the higher the influence to a factor is assigned. OAT influences are often expressed as probabilities or frequencies, which are hard to interpret due to rare event characteristics of accidents, while uncertainty-based concepts like likelihoods are rarely used because of difficulties with qualitative interpretation of likelihood values, (Trucco, et al., 2008), (Chen, et al., 2019). Moreover, such measures do not deliver the information on whether the influence factor contributes favourably or unfavourably to the accident occurrence; nor do they provide any qualitative scores. Qualitative scores used for interpretation of influence factors in maritime domain have been previously addressed in (Mazaheri, et al., 2016), using subjective, expert elicited weights based on the uncertainty of experts' knowledge. We are not aware of other advancements along this line of research, even though the availability of objective qualitative interpretation is important for a wider acceptance and practical deployment of the probabilistic analysis of maritime accidents.

In our paper, we have introduced the weight of evidence (WoE) approach, (Good, 1985), (Osteyee & Good, 1974), a likelihood-based approach, and we have derived the strength of evidence (SoE) as a quantitative measure that enables an interpretation of influence factors by means of qualitative categories, inspired by Bayes Factor, which are easily comprehensible to users of different backgrounds. The introduction of WoE and SoE is made possible by conceptualizing the Bayesian network model as a Bayesian network classifier. Bayesian Network classifier, (Friedman, et al., 1997), (Chan & Darwiche, 2002), (Bielza & Larrañaga, 2014), represents a decision function that distinguishes accident influencing factor state spaces into those which contribute to an increased chance of accident and those which contribute to a reduction of the chance of accident. This conceptualization importantly differs Bayesian Network classifier model from the common practice Bayesian Network models which do not take into account the non-accident scenarios and thus cannot assess model factors discriminatively with respect to their dual influence on the outcome.

This paper contributes to a current probabilistic maritime accident modelling and assessment methodology by defining the maritime accident model as a binary classifier, by introducing the likelihood based inference measure, and providing the

grading scale for influence factor interpretation. The overview of the complete framework for risk assessment proposed in this paper is summarized in Fig.1. In the first phase, the BN classifier is structured and parametrized based on an expert elicitation and available data records. Credibility assessment is performed to verify the behaviour of the BN classifier. In the second phase inference is performed based on developed strength of evidence measure. Accordingly, the rest of the paper is organized as follows: in Section 2. a short theoretical background of Bayesian network is introduced, and accident formulation as a binary classification problem is defined. Next, the weight of evidence approach is developed. In section 3. the explication of model development and credibility assessment are given. The results of the approach are demonstrated in Section 4. The concluding remarks are given in Section 5.

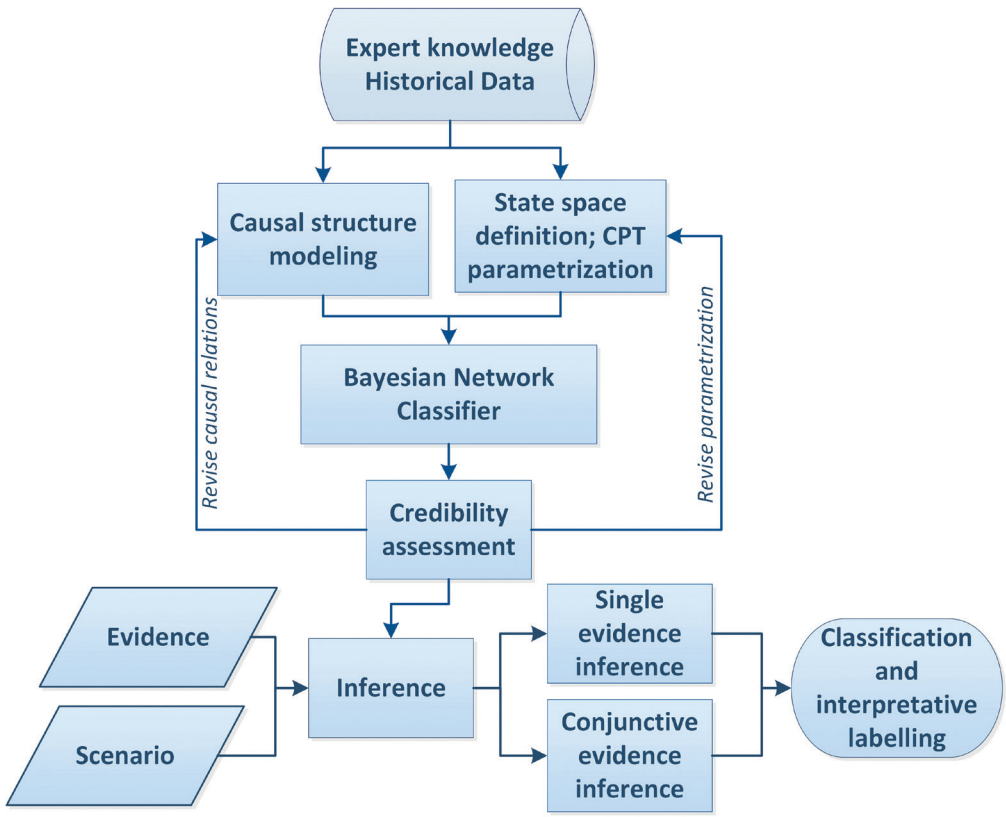


Figure 1.
The framework for influence factor assessment based on Bayesian Network classifier.

2. BAYESIAN NETWORK

Bayesian network, (Pearl, 1988), (Jensen & Nielsen, 2007), (Darwiche, 2014), offers a unified modelling framework that compensates for insufficiency of statistical information and uncertainty of expert knowledge and is thus able to encode sparse data and different aspects of experts' knowledge and beliefs about the maritime accidents. Bayesian network is formally defined as triplet (V, G, Θ). V denotes a set of n random variables $V = \{V_1, V_2 \ldots, V_N\}$, G is a directed acyclic graph whose nodes are members of V connected in such a way that each variable is conditionally independent of its non-descendants given its parents. Each directed edge represents a conditional dependence between parent-child node pair. Let the parents of $Vi \in V$ in G be denoted by $\pi(Vi)$; let Θ denote the set of local conditional probability distributions $\Theta = \{P(V_i \mid \pi(V_i)), V_i \in V\}$. Given (V, G, Θ), BN provides a joint probability distribution over the set of V as multiplication of a set of conditional probabilities:

$$P(V) = \prod_{V_i \in V} P(V_i \mid \pi(V_i)) \quad (1)$$

A set of variables V is organized into three types of nodes: observable nodes, intermediary node, and target nodes, while variable dependencies are described with three causal classes: causal chain, common effect, and common cause, (Pearl, 2000). Observable nodes are nodes for which statistical data or strong knowledge is available, such as weather conditions, availability of technical equipment and similar. Intermediary nodes are unobservable or partly observable nodes, such as human factors, human related factors, technical related factors, etc., for which the experts' beliefs or limited historic data is available. The observable and intermediary nodes both form the set of influence nodes, and will be denoted by $X = \{ X_1, X_2, ..., X_n \}$, $n = N - 1, X \subset V$. The remaining variable from the set $V \setminus X$ is a target node, representing maritime accident decision node, and will be denoted as Y throughout the paper.

The process of development of Bayesian network is organized into tasks: domain knowledge is acquired, relevant hazard factors to constitute the set of variables V are identified and causally connected, and probabilities and conditional probabilities are assigned. An iterative process is required during the model development. Multiple revisions of associations that form the G and check of values in Θ are performed. Both qualitative and quantitative BN development follows good practice guidelines for BNs in safety and reliability analysis domains. The interested reader is referred to (Chen & Pollino, 2012), (Marcot & Penman, 2019), (Sigurdsson, et al., 2001).

2.1. Accident Formulation as a Binary Classification and Context Definition

Maritime accident BN models are typically built upon data available in accident records and data elicited from human expert knowledge of accident scenario. A drawback of approaches based on accident records and knowledge of accidents is that they do not take into account non-accidents, because there is not such data, (Stornes, 2015), and thus cannot provide sufficient confidence in estimating influencing factors. This point has been discussed concisely from a methodological viewpoint in (Øvergård,2015) In our paper we take on an approach that seeks to include the non-accident data as well through expert elicited data on non-accident. Therefore, we pose the accident modelling as a binary classification problem where two targeted states are "accident occurring" and "accident non-occurring". A modification of common approach is made at the point of data collection from experts, where the expert knowledge elicitation is made for accident and non-accident cases.

Since the BN framework supports the definition of probabilities as degrees of belief, the probabilities represented as the degree of belief are used to define probabilities of events that occur rarely or have not yet occurred. This way, non-accidents, for which no real-world data exists, can be defined by experts. Often applied frequentists definition of probability requires an event to have occurred enough times to allow the collection of accountable information which in this case is not possible as maritime accident is a rare event.

We seek the approach, the one that exploits the fact that experts can define non-accidents scenario and propose beliefs as inputs to BN accordingly. In our work, the risk assessment model reflects current navigational situation scenario in which a maritime accident is a rare event, thus the focus is on the modelling of the current state-of-the-environment of the small ship in navigation under collision risk. The model is conceptualized in such a way as to include all factors believed to have the ability to cause an accident, but also to reduce the chance of accident; and it is parametrized in such a way as to include both data, where available, and expert knowledge of the accident and expert knowledge and belief for non-accident scenarios.

Variables $X = \{ X_1, X_2, \ldots, X_n \}$ and variable states x^i_j, $i \in [1,..,n]$, $j \in [2,..,m]$, known as influence factors, encode structurally and parametrically the likelihood of occurrence or non-occurrence of an accident Y in accordance to data and expert elicitation. The accident variable Y has two states, y and $\bar{y}$, that correspond to "accident" and "non-accident", respectively. Whenever a random variable X takes on a state value $x = e$, it is called an evidence. Given a BN, for which the prior probability distribution of the target node defines the probability threshold $h_0 = \frac{P(y = y_0)}{P(\bar{y} = \bar{y}_0)}$, there exists a classification function $F^{BN}(x)$ that assigns labels $\{0, 1\}$ to influence factors x by evaluating likelihood of accident occurrence given the evidence, based on the weight of evidence, (Osteyee & Good, 1974), as follows:

$$F^{BN}(x=e) = \begin{cases} 0 \text{ if } \frac{P(y|x=e)}{P(\bar{y}|x=e)} > h_0 \\ 1 \text{ if } \frac{P(y|x=e)}{P(\bar{y}|x=e)} < h_0 \end{cases} \quad (2)$$

The classification labels have a semantic interpretation. A factor x is "accident contributing" or has an "adverse influence" if $F^{BN}(x) = 0$, or factor x is "accident preventing", or has a "beneficial influence", if $F^{BN}(x) = 1$. When $\frac{P(y|x=e)}{P(\bar{y}|x=e)} = h_0$ the factor is on a decision boundary and its individual influence is neutral. Influence factors, labelled 0, are causative and will be denoted x, while those labelled 1, are preventive, and will be denoted $\bar{x}$. Classifier concept enables reasoning with BN, and it is extended further by the analysis of weight of evidence to allow grading of $P(y|x=e)$ and $P(\bar{y}|x=e)$ according to the strength of the response of the target variable. Details are developed in the Subsection 2.2.

Maritime collision context definition

The development of BN classifier that would generalize over a spectrum of environments would not be feasible due to specificity and uncertainty inherent to diversity of possible scenarios; therefore, it is required to constrain the context of risks assessment in terms of the accident type and the environment. The model development and the approach to risk assessment is exemplified on maritime collision as an accident type being of the most interest to Croatian nautical tourism safety assessment due to high severity of consequences, though the approach and methodology can be generalized to other maritime accident types. Of interest in this paper are small passenger ships in non-linear coastal navigation, having a length of below 70 meters with maximum allowed capacity of 250 persons. The safety risk of these vessel types carrying passengers on a commercial basis might occur in particular during the tourist season, when the density of the sea traffic is considerably increasing. Safety concerns are further accentuated by the fact that coexisting risk factors, such as technical, human, and environmental parameters could possibly contribute to an unfavourable event, collision being the most severe in terms of harm for human lives and assets.

2.2. Weight of Evidence Approach to Reasoning With Bayesian Network Classifier and Interpretation

Not every evidence *x* has an equally strong impact on the target node. The strength of the impact of evidence on the target can be quantitatively measured by the adaptation of Irvine J. Good's weight of evidence approach, (Good, 1985), (Osteyee & Good, 1974). The discriminative model formulation allows, not only for classification, but for testing of the strength of evidence with respect to two competing hypotheses, H_1 and H_2.

The weight of evidence is the difference in information about *x* provided by H_1 compared to H_2, (Osteyee & Good, 1974):

$$WoE(H_1/H_2:x) = I(x:H_1) - I(x:H_2) \tag{3}$$

$$WoE(H_1/H_2:x) = \log\frac{P(x|H_1)}{P(x)} - \log\frac{P(x|H_2)}{P(x)} = \log\frac{P(x|H_1)}{P(x|H_2)}$$

The weight of evidence, as the log likelihood of the evidence given the two hypotheses, is further developed according to the Bayes theorem:

$$WoE(H_1/H_2:x) = \log\frac{P(H_1|x)}{P(H_2|x)} - \log\frac{P(H_1)}{P(H_2)} \tag{4}$$

$WoE(H_1 \setminus H_2 : x)$ in Eq.(3) can be positive or negative. The positive *WoE* means that hypothesis H_1 is supported by *x*, while negative *WoE* indicates support of *x* to H_2. Classifier function in Eq. (2) evaluates these properties of Eq.(4) and assigns labels to evidence accordingly. Besides classification, we are interested in measuring the strength with which the evidence *x* contributes to the hypothesis. Now, we take that H_1 is an accident (or non-accident) hypothesis, which competes with the baseline case hypothesis, H_2. In this case, it is taken that prior probabilities of the two hypotheses are equally probable, which eliminates the second term of Eq. (4). Therefore, the strength of evidence for H_1 against H_2 is measured as the absolute value of the first term in Eq. (4):

$$SoE(H_1/H_2:x) = \left|\log\frac{P(H_1|x)}{P(H_2|x)}\right| \tag{5}$$

The strength of evidence as the log ratio of $P(H_1|x)$ and $P(H_2|x)$ is a measure of relative change. Since $P(H_1|x)$ and $P(H_2|x)$ are close values (relatively small changes are expected), the log ratio of the two conditional probabilities can be approximated with a percent change. Therefore, we define $SoE(H_1 \setminus H_2 : x)$ as the percent change as follows:

$$\%SoE(H_1/H_2:x) = \left|\frac{P(H_1|x) - P(H_2|x)}{P(H_2|x)}\right| \tag{6}$$

Causal reasoning with probabilistic models depends on interpretations of the results of WoE by means of $F^{BN}(x=e)$ and $SoE(H_1 \setminus H_2 : x)$. Standard grading of evidence into categories used for interpretation of the *WoE* are derived from Bayes Factor (*BF*) scales, and are transformed as 10log (*BF*), (Jeffreys, 1998), (Kass & Raftery, 1995). Absolute values of *BF* and its interpretations are shown in Table 1. Adaptations of scales of *WoE* are not uncommon, (Kass & Raftery, 1995). In our paper the modification is made to the first two grading categories whose role in general applications is to eliminate irrelevant evidence. To accommodate the expert elicitation and Bayes network classifier parametrization process, through which insensitive and irrelevant factors have already been eliminated, the ranges of the first two categories have been adapted to the application, and therefore changed from (0 to 5) and (5 to 10) to [0,1 ⟩ and [1,10 ⟩. The proposed interpretation of the importance of influence factors for the classification system in our paper is given in Table 2.

Using the interpretative categories proposed in Table 2, influence factors, either causative, x, or preventive, $\overline{x}$, can be verbally labelled. Causative influence factors are interpreted with respect to their influence to cause the accident, while preventive influence factors are interpreted with respect to their potential to prevent accident.

Table 1.
Linguistic interpretation of value categories of the weight of evidence (from (Jeffreys, 1998), (Kass & Raftery, 1995)).

Evidence against the null hypothesis	Bayes factor (BF)	WoE=10 log(BF)
Anecdotal evidence	1 to 3.2	0 to 5
Substantial evidence	3.2 to 10	5 to 10
Strong evidence	10 to 100	10 to 20
Decisive evidence	>100	>20

Table 2.
Interpretative categories of influence factor based on percent change obtained by %SoE.

Classification of influence factor	Interpretation of relative influence on the target variable (accident/non-accident)	SoE [%]
weak influence factor	its influence on the target variable is not critical	0 - 0.9
substantial influence factor	its influence on the target variable is significant	1 - 9.9
strong influence factor	its influence on the target variable is very significant	10 - 19.9
extremely strong influence factor	its influence on the target variable is critical	>20

Using the interpretative categories proposed in Table 2, influence factors, either causative, x, or preventive, $\overline{x}$, can be verbally labelled. Causative influence factors are interpreted with respect to their influence to cause the accident, while preventive influence factors are interpreted with respect to their potential to prevent the accident.

3. STRUCTURING AND PARAMETRIZATION OF THE BASELINE BAYESIAN NETWORK CLASSIFIER FOR COLLISION

A first line in BN construction is the definition of general hazard types for maritime accidents and their relevant factors. To identify general hazards for collision, a review of the existing knowledge in literature has been made, while institutional database information has been collected and interviews with maritime experts have been held in order to define specific local hazard conditions, extract relevant factor states, and establish causal relations among factors. Hazard factors, in maritime accidents, may be directly observable (weather, equipment, wind), but also many factors are unobservable due to their intrinsically unobservable nature (e.g. human error, personal condition, etc.) or the lack of indications of presence or absence of such factors. Some factors may be partially observable (e.g. safety culture). These factors are intermediary nodes of the causal network. Relevant factors are identified with the help of existing literature on maritime accidents from (Hänninen & Kujala, 2012) and (Mazaheri, et al., 2016), and adapted for collision context analysed in this paper. The compilation of identified factors helps the construction of causal relations and the formation of the causal system encoded with the network topology. In our approach, the network topology growth is initialized at target node 'Collison' and parental nodes are added using the rules of the previously mentioned three causal classes, where interpretation of direction of arrow from node Vi to Vj is that belief in Vi implies expectation in Vj, (Jensen & Nielsen, 2007). Addition of the first level causal nodes: "Give-way", "Communication with other ship", "Loss of control of other ship" and "Traffic distribution" is based on a minimal theory-based model deduced from COLREG rules, (Cockcroft & Lameijer, 2003), and from collision avoidance strategies in an interaction during a critical encounter situation, (Chauvin, et al., 2013), (Chauvin & Lardjane, 2008). Next, each of the first-level parental nodes are further explained and related to their own parental nodes, thus embedding further influence factors from the preselected factor list. In a similar manner, the network is grown until observable independent factors are reached.

It should be noted that development involves refinement of structure through adding the omitted factors and removing irrelevant factors and their relations through collaboration with domain experts. Fig. 2. shows resulting causal influence network of the maritime collision, its nodes and dependence structure. In

the subsequent step, state spaces of each factor are defined. For example, tiredness is an influence factor, and it has two states: present or not present; availability or nonavailability of the technical equipment are separate aspects for technical factors; wind force has multiple states, etc. Next, conditional probability distributions of each factor given its parents are assessed and its adjustments are performed, in a top-down manner, from the causal nodes to the effect node. Conditional probabilities are assigned based on expert's beliefs, national weather reports, local maritime databases, and the existing literature. The final parametrisation of the state space is based on a consensus elicitation of expert knowledge, (Hassall, et al., 2019), (Zhang & Thai, 2016). The description of each factor, their state spaces and sources of data, are given in the Appendix.

Local coherence checking is conducted during this stepwise parametrisation of nodes using sensitivity analysis, (Kjærulff & Gaag, 2000), implemented in GeNIe/SMILE, (BayesFusion, 2020), and the parameter calibrations are performed accordingly. The model encapsulates the collective domain knowledge of the expert group participants and their understanding of safety problem of the maritime collision risk, thereby representing the baseline model.

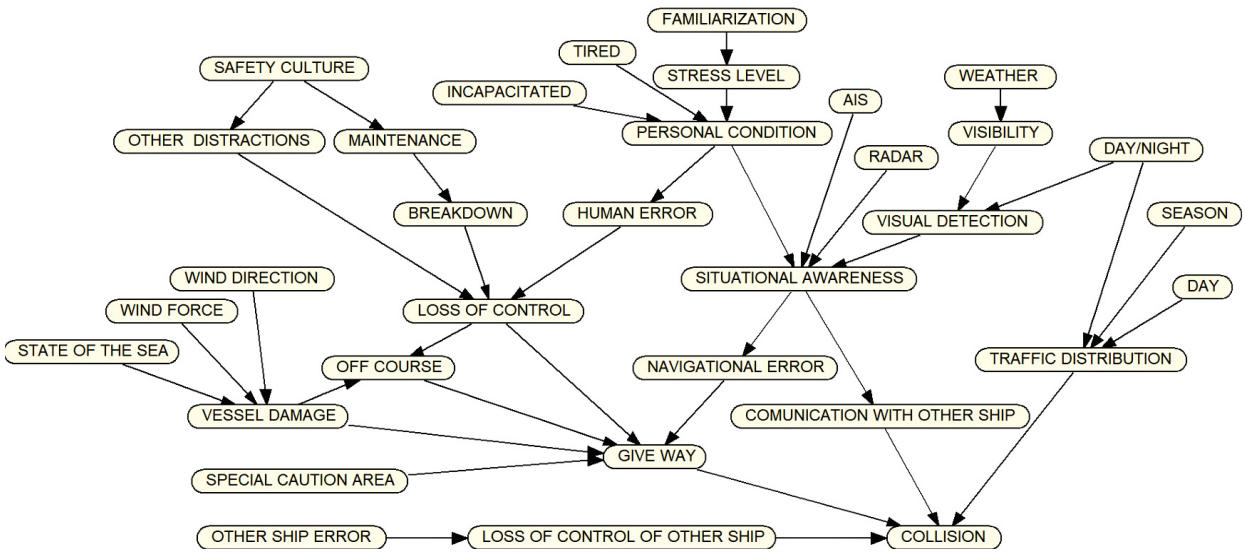


Figure 2.
The causal influence network for maritime collision.

An individual ship involvement in a collision accident is modelled, as such models are underrepresented in scientific literature, (Ozturk & Cicek, 2019). Most of the proposed collision models are built from the point of view of two ships involved. While increasing the complexity of the Bayesian Network and accompanying computational complexity, such models do not yield the gain in information on influential risk factors for individual ship, as they do not address interactions of variables of complex and large Bayesian Networks. We believe that the behaviour of individual ship should be well developed and understood prior to modelling a collision as a two-party event. It should be noted that in our model an influence of another ship is not neglected, yet its contribution is represented as a simplified sub-network comprised of chain "Other ship error" → "Loss of control".

3.1. Credibility of the Baseline Bayesian Network Classifier Model

A common final stage of any classification model development, before its deployment, is the assessment of its accuracy based on the real-world data. When data-driven assessment is not feasible, i.e., when there is no data to confront the developed model with the credibility of the model is assessed. Credibility verifies that underlying assumptions and properties of the model are satisfied. Monotonicity is the property that should be exhibited by expert knowledge based Bayesian networks, thus, a way to evaluate credibility of the developed BN classifier model is the verification of monotonicity, (Pianosi, et al., 2014), (Gaag, et al., 2004). Monotonic behaviour is incorporated into a model through qualitative influences at the development

stage. Violation of monotonicity can occur despite engineering efforts to carefully encode knowledge and it can exist even after sensitivity tests were performed, (Plajner & Vomlel, 2017), (Pianosi, et al., 2014). A fundamental monotonic behaviour is observed when the increase in values of parental variables produce the increase in corresponding values for the child variables. For example, let X_i be the parent of Y. Let the parental influence factors x and $\bar{x}$ be the causative and preventive, respectively; and let child nodes' states y and $\bar{y}$ correspond to "accident" and "non-accident". Whenever y takes on a value 1, it represents an evidence that accident occurs, and at the same time $\bar{y}$ becomes 0. Let this be denoted as Y_{acc}. Likewise, whenever $\bar{y}$ takes on a value 1, it represents an evidence that accident does not occur, and at the same time y becomes 0. Let this be denoted as $Y_{non\text{-}acc}$. The exemplary parental conditional probabilities are given in Table 1. Due to monotonicity constraint, it should hold for every parental node that $P(x \mid Y_{acc}) \geq P(\bar{x} \mid Y_{acc})$ and $P(\bar{x} \mid Y_{non\text{-}acc}) \geq P(x \mid Y_{non\text{-}acc})$. Any parent-child nodes (i.e. effect-cause nodes) should respect these relations locally. Using a bottom up approach, from the target node backward to root nodes, the whole network is checked locally for monotonic behaviour.

Table 3.
Example of the conditional probability table of parent-child nodes.

	collision :: yes	collision :: no
influence factor :: preventive	$P(\bar{x} \mid Y_{acc})$	$P(\bar{x} \mid Y_{non\text{-}acc})$
influence factor :: causative	$P(x \mid Y_{acc})$	$P(x \mid Y_{non\text{-}acc})$

Also, the global monotonicity test can be performed with monotone likelihood ratio test (MLR), (Mukhopadhyay, 2000), a measure similar to odds ratio, which, unlike the odds ratio, assumes a causal relation of the influence factor and the outcome:

$$MLR = \frac{\frac{P(\bar{x} \mid Y_{acc})}{P(\bar{x} \mid Y_{acc}) + P(\bar{x} \mid Y_{non\text{-}acc})}}{\frac{P(x \mid Y_{acc})}{P(x \mid Y_{acc}) + P(x \mid Y_{non\text{-}acc})}} \tag{7}$$

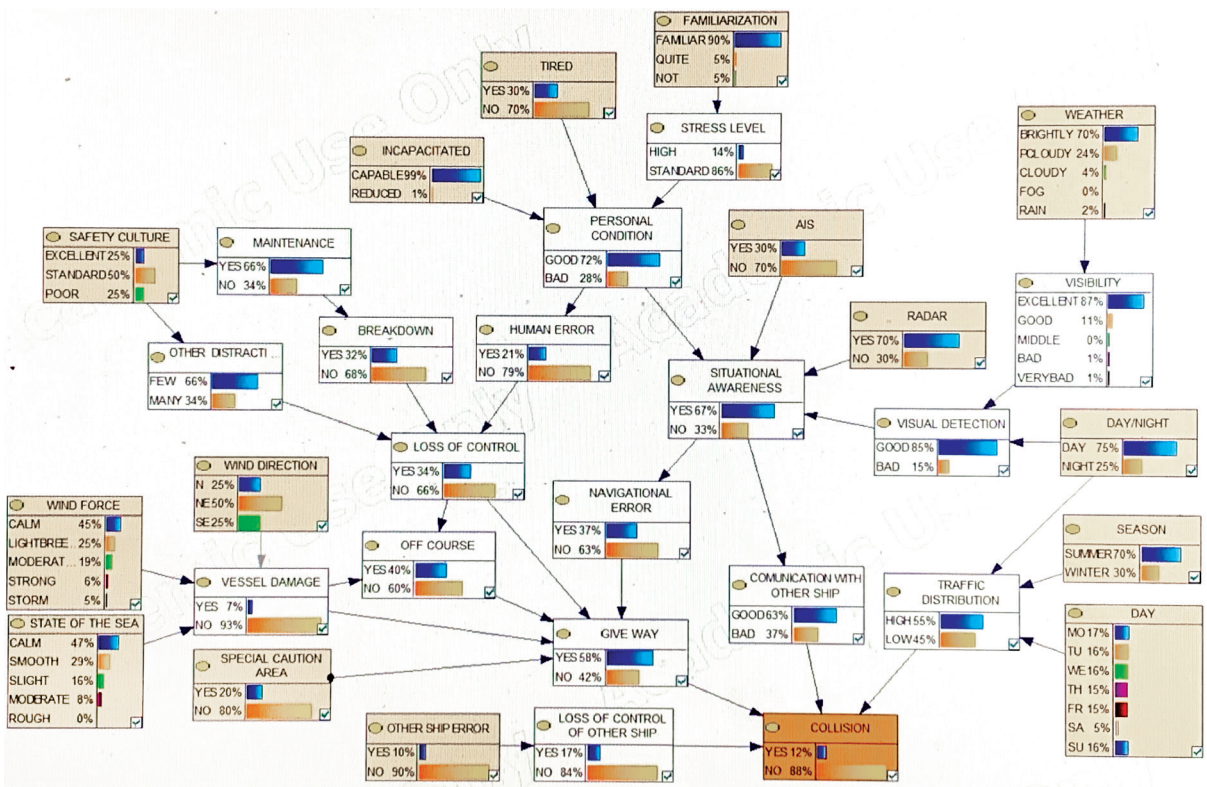


Figure 3.
Baseline model inference (developed and tested using GeNIe/SMILE, (BayesFusion, 2020) and pyAgrum, (Gonzales, et al., 2017)).

MLR indicates how much more likely is it that a collision will occur when preventive influence factor is present compared to when causative influence factor is present. Due to monotonicity, it should hold that MLR ≤ 1, which indicates that when a random variable takes a preventive state, the likelihood of accident is lower than (or equal to) the likelihood of accident when the same random variable takes on a causative state. When any of tested links does not obey monotonicity suggested by expert, it informs us that an intervention in conditional probability tables is required. In this case, non-monotonic influence factor should be optimized and recalibrated using the sensitivity analysis.

The inference results of credibility tested baseline Bayesian Network classifier show that the likelihood of collision occurrence is 12%, compared to the likelihood of 88%, that collision will not occur under baseline state of the world situation (Fig.3.). This implicates that substantial belief is held by experts against the collision occurrence in the baseline situation, given encoded uncertainty about the state of the world. The GeNIe/SMILE, (BayesFusion, 2020), is used for the development and Phyton library pyAgrum, (Gonzales, et al., 2017), for inferences. Belief updates in the network are based on Lazy propagation algorithm, (Madsen & Jensen, 1999).

4. RESULTS OF INFERENCE WITH BAYESIAN NETWORK CLASSIFIER

Evidence analysis based on the weight of evidence approach delivers two pieces of information: classification of causative and preventive influence factors, and their strength of influence (SoE) with interpretative categories. Below, the results of single evidence inference analysis and conjunctive evidence inference analysis will be presented.

Single evidence inference results are presented in Table. 4. Influences based on the SoE metric and the colour coding refer to the interpretation system proposed in Table 2. Collision risks are identified as collision contributing, adversely influencing factors, termed causative influence factors. The proposed method identifies accident preventing contributors as well. The left-side column presents the strengths of causative influence factors, and the right-side column presents the influence factors that have a collision preventing potential.

Several extremely strong single risk factors identified are poor communication with other vessel, lack of situational awareness, loss of control of other ship, failing to comply with a give way, human error, loss of control, and poor personal condition. According to the SoE scoring, poor communication with other vessel raises the relative likelihood of collision by 28.6%, and its influence is estimated as a critical risk. The communication between vessels has been identified as one of the most influential factors in other studies of risks in navigational situations involving small passenger vessels and pleasure craft in high density area, (Øvergård, et al., 2020). The lack of situational awareness raises a relative likelihood of collision by 25.7% and is assessed as critical risk. This is in accordance with common maritime safety knowledge and many documented accidents. USCG accident database indicates that the lack of situational awareness is the causal factor in 60% of all accident causes, (Baker & McCafferty, 2005). Next, the factor contributing critically to accidence occurrence is a poor personal condition, which is in direct causal relation to situational awareness and an immediately preceding parent of human error. Needless to accentuate, human related factors are credited as the major cause of accidents. Among strong preventive influences, compliance with give way regulations and the communication with other vessel are sharing the first place, followed by low traffic density, no navigational error, staying in control of vessel, winter season and being on the navigational course. In the preventive factors class list of Table 4, no extremely strong preventive influence factors are identified, which is the expected result of the classifier, considering that it has been organized and developed to reflect the normal collision-free state of navigation under the risk of collision.

In the interpretation of BN classifier inference results, one should be aware that evidence influence measured by the strength of evidence might not be the same as the importance of evidence. Namely, in the interpretation of inference scores provided by the strength of evidence, the proximity of the evidence node to the target node can be considered. It is known that nodes proximal to the target node of causally conceptualized networks are more strongly affecting the target node, (Hänninen & Kujala, 2012), and the influence of evidence on the target node attenuates with the propagation length between these nodes, (Yuan & Druzdzel, 2007). A simple propagation length weighting, that balances out the influence of adjacency of nodes, is introduced to perform ranking of influence factors, where propagation length is measured with the depth of graph, (Yuan & Druzdzel, 2007). Propagation length weight we is defined as the ratio of minimal number of edges traversed from evidence to target node, and maximal number of edges traversed from the deepest node. In our model, the maximal depth is 7, as counted from "Familiarization" to "Collision". So, for example, for the "Human error" variable, the propagation length weight is $w_e = 3/7$.

Table 4.
Classification of causative and preventive influence factors, their influence strengths and interpretative categories (red: extremely strong influence; orange: strong influence; green: substantial influence; yellow: weak influence. See Table 2.)

causative influence factors	SoE (%)	preventive influence factors	SoE (%)
COMUNICATION WITH OTHER VESSEL::POOR	28.6322	GIVE WAY::YES	16.6576
SITUATIONAL AWARENESS::NO	25.7142	COMMUNICATION WITH OTHER VESSEL::GOOD	16.6576
LOSS OF CONTROL OF OTHER SHIP::YES	24.5791	TRAFFIC DENSITY.::LOW	13.0642
GIVE WAY::NO	23.2998	SITUATIONAL AWARENESS::YES	12.9394
HUMAN ERROR::YES	21.0030	NAVIGATIONAL ERROR::NO	11.0112
LOSS OF CONTROL::YES	20.7900	LOSS OF CONTROL::NO	10.8181
PERSONAL CONDITION::POOR	20.4178	SEASON::WINTER	10.7047
NAVIGATIONAL ERROR::YES	18.9269	OFF COURSE::NO	10.3557
INCAPACITATED::REDUCED	17.5977	DAY/NIGHT TIME::NIGHT	7.8712
OTHER SHIP ERROR::YES	17.2202	PERSONAL CONDITION::GOOD	7.7834
OFF COURSE::YES	15.4193	HUMAN ERROR::NO	6.2625
VESSEL DAMAGE::YES	15.2623	LOSS OF CONTROL OF OTHER SHIP::NO	4.8567
TRAFFIC DENSITY::HIGH	10.6247	TIRED::NO	3.1742
WIND FORCE::STORM	9.2608	AIS::YES	2.1032
STRESS LEVEL::HIGH	8.5240	RADAR::YES	2.0824
TIRED::YES	7.4072	DAY::{FRI,SAT}	1.9855
FAMILIARIZATION::NOT	6.0461	OTHER SHIP ERROR::NO	1.9131
RADAR::NO	4.8596	OTHER DISTRACTIONS::FEW	1.604
SEASON::SUMMER	4.5881	SAFETY CULTURE::EXCELLENT	1.5752
VISIBILITY::VERY POOR	4.3718	BREAKDOWN::NO	1.5287
WEATHER::RAIN	4.0371	STRESS LEVEL::STANDARD	1.3874
BREAKDOWN::YES	3.2586	MAINTENANCE::YES	1.26
OTHER DISTRACTIONS::MANY	3.1493	VESSEL DAMAGE::NO	1.0682
SAFETY CULTURE::POOR	2.7765	WIND FORCE::CALM	1.0682
DAY/NIGHT TIME::DAY	2.6240	VISUAL DETECTION::GOOD	0.4173
MAINTENANCE::NO	2.4740	FAMILIARIZATION::FAMILIAR	0.3962
VISUAL DETECTION::POOR	2.3348	SPECIAL CAUTION AREA::NO	0.3902
SPECIAL CAUTION AREA::YES	1.5621	INCAPACITATED::CAPABLE	0.1775
AIS::NO	0.9017	SEA STATE::CALM	0.17
SEA STATE::MODERATE	0.8098	VISIBILITY::EXCELLENT	0.0822
DAY::{MO,TU,WE,TH,SU}	0.4967	WEATHER::BRIGHT	0.0822
WIND DIRECTION::NE	0.0002	WIND DIRECTION::SE	0.0002

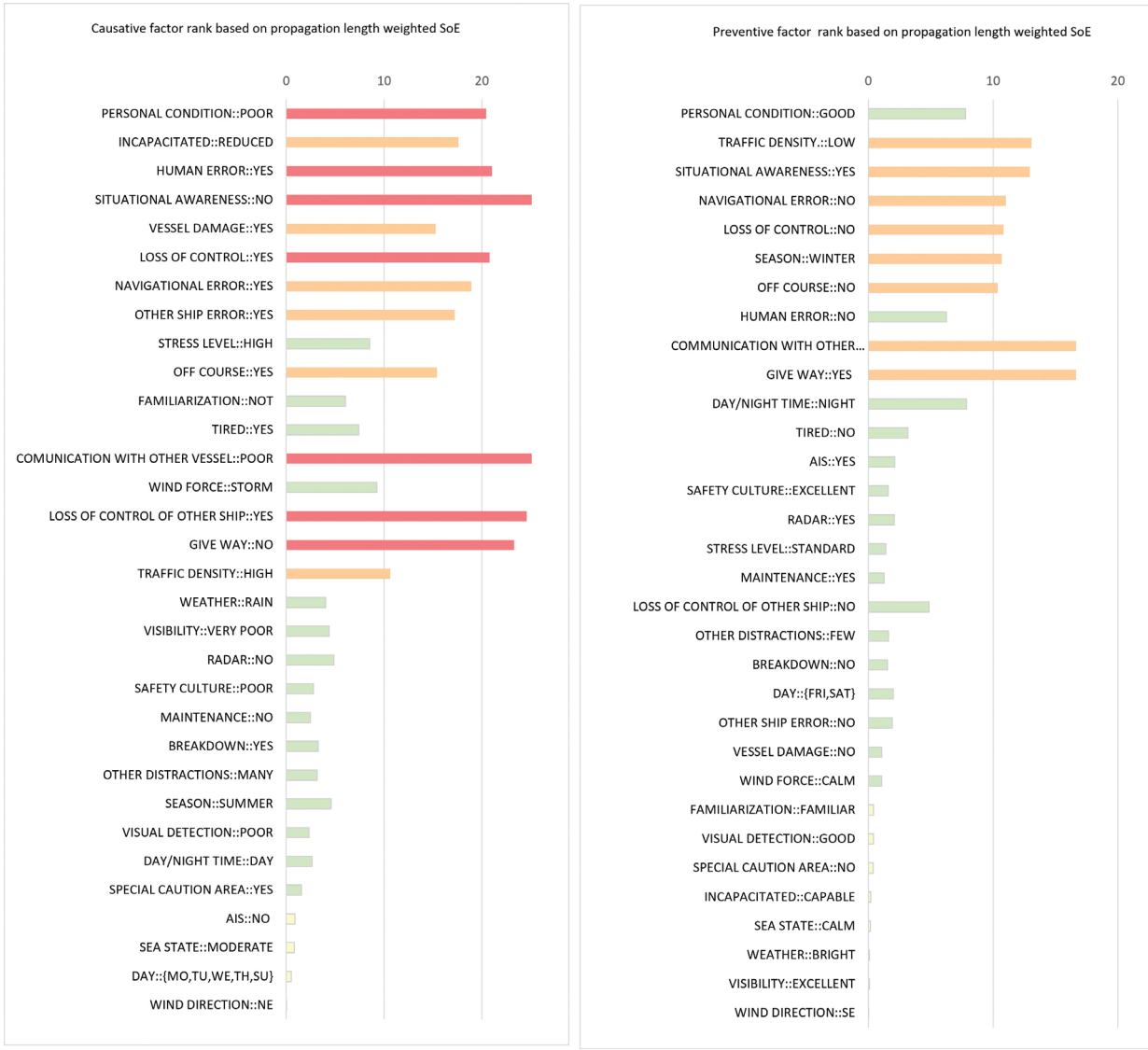


Figure 4.
Importance ranking based on propagation length weighted SoE.

According to the ranking based on weighted SoE, extremely strong and the strong causative, as well as the strong preventive influencing factors, are obtained and shown in Fig 4. It yields that the extremely strong causative influencing non-adjacent factors are poor personal condition, human error, lack of situational awareness, and a loss of control. This reasoning is supported by the results of preventive factor ranking, where good situational awareness is the highest ranked preventive factor. Previous studies and experience have recognized contribution of these human related factors and a lack of situational awareness in the collision causation, and in particular in the collision causation of small passenger vessels that are not obligated to carry a radar and/or AIS.

The influences of the observable risks and preventive factors, often called indicators, for an individual ship are extracted in Table 5. Among observable factors, strong risk for collision is a reduced psychophysical condition of the person responsible for watchkeeping. This is embodied in an 'incapacitated::reduced' factor whose name is adopted from (Hänninen & Kujala, 2012). With regard to assessment of influences of technical equipment, the results show that not being equipped with a radar presents a higher risk of collision that not being equipped with AIS. Not

having a radar increases relative risk of collision by almost 5%, and not being equipped with AIS raises relative risk by nearly 1%. Also, the results show that being equipped with a radar or AIS, individually contribute preventively by reducing the likelihood of collision by around 2% (factor influences are not necessarily additive). Among environmental factors, the summer season is ranked as a substantial causative factor, while the winter season is ranked as a factor with strong preventing potential. This implies the importance of the investigation of traffic density which is in direct relation to season.

Table 5.
Classification of causative and preventive observable influence factors and SoE categories.

observable risk factors	SoE (%)	observable preventive factors	SoE (%)
INCAPACITATED::REDUCED	17.5977	SEASON::WINTER	10.7047
WIND FORCE::STORM	9.2608	DAY/NIGHT TIME::NIGHT	7.8712
TIRED::YES	7.4072	TIRED::NO	3.1742
FAMILARIZATION::NOT	6.0461	AIS::YES	2.1032
RADAR::NO	4.8596	RADAR::YES	2.0824
SEASON::SUMMER	4.5881	DAY::{FRI,SAT}	1.9855
WEATHER::RAIN	4.0371	OTHER SHIP ERROR::NO	1.9131
SAFETY CULTURE::POOR	2.7765	SAFETY CULTURE::EXCELLENT	1.5752
DAY/NIGHT TIME::DAY	2.6240	STRESS LEVEL::STANDARD	1.3874
SPECIAL CAUTION AREA::YES	1.5621	WIND FORCE::CALM	1.0682
AIS::NO	0.9017	FAMILIARIZATION::FAMIL.	0.3962
SEA STATE::MODERATE	0.8098	SPECIAL CAUTION AREA::NO	0.3902
DAY::{MO,TU,WE,TH,SU}	0.4967	INCAPACITATED::CAPABLE	0.1775
WIND DIRECTION::NE	0.0002	WEATHER::BRIGHT	0.0822

The great increase in traffic density during the tourist season has raised particular concerns about the navigational safety of small tourist passenger vessels, and the assessment of the influence of traffic density on the collision causation is one of the main interests that have spurred the risk assessment approach development described in this work. According to the results, the high traffic density rises the relative likelihood of collision by 10%. Additional confirmation that traffic density in general is a strong factor is given by the observation in the results that the low traffic density has a preventive potential. Therefore, to gain a more in-depth understanding of the influence of high traffic density, a conjunctive evidence analysis is performed, with an aim to identify the co-occurring risks. The influence of the conjunction of high traffic density and all other factors on accident occurrence is given in Table.6.

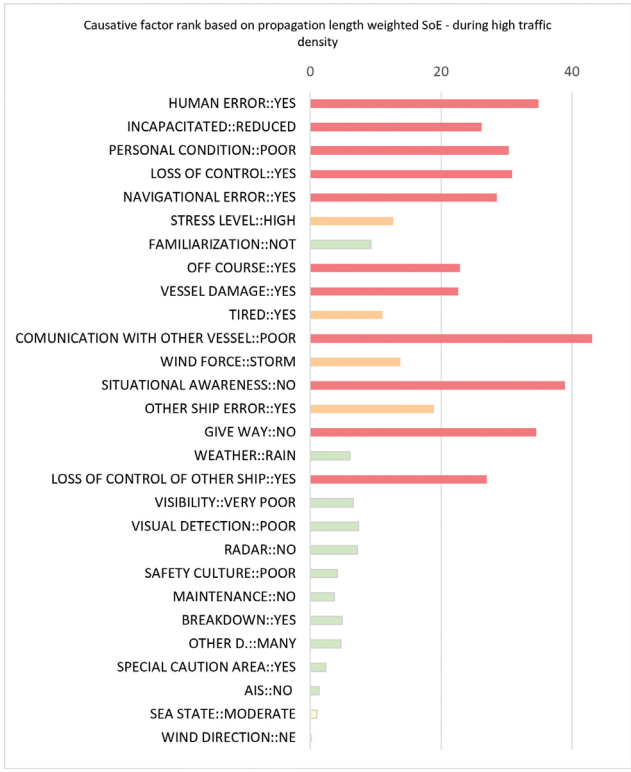
Conjunctive evidence analysis results reveal that during the high maritime traffic density, the number and the strength of co-occurring causative influence factors have risen, while the number of preventive factors has been reduced, but also strengthened. Human error is a novel extremely strong risk factor when co-occurring during the high traffic density, as can be revealed through comparison of difference of single risks factors and those risks factors arising through co-occurrence analysis (from comparison of data between Table 4 and Table 6). Extremely strong factors, both causative and preventive, are the communication with other vessels and give way, which accentuates the important influential character of these factors, preceding immediately the occurrence of collision. Among observable risk factor, incapacitation, i.e. reduced psychophysical ability, is heightened, and its influence is interpreted as an extremely strong collision risk (Table 6). Also, stormy wind, often occurring suddenly during high traffic tourist season, and tiredness become very significant risk factors. No observable preventive factor can be extracted, which means that prevention cannot be focused on any indicator factor but on an unobservable, intermediary once. Again, to compensate for adjacency of these nodes to targeted collision node, the ranking is performed based on propagation length weighted SoE, and shown in Fig. 5. The propagation length weighted SoE reveals that human related factors are most prominent collision contributing factors. Among preventive factors, the good situational awareness is accentuated as a very significant preventive factor during high traffic density.

Table 6.
Classification of causative and preventive influence factors and their SoE based interpretative categories for the assessment of collision during high maritime traffic density period.

causative influence factors	SoE (%)	preventive influence factors	SoE (%)
COMUNICATION WITH OTHER VESSEL::POOR	43.0723	GIVE WAY::YES	24.642
SITUATIONAL AWARENESS::NO	38.9043	COMUNICATION WITH OTHER VESSEL::GOOD	24.642
HUMAN ERROR::YES	34.89	SITUATIONAL AWARENESS::YES	19.1552
GIVE WAY::NO	34.5448	NAVIGATIONAL E.::NO	16.312
LOSS OF CONTROL::YES	30.8289	PERSONAL CONDITION::GOOD	11.5343
PERSONAL CONDITION::POOR	30.3519	OFF COURSE::NO	6.3415
NAVIGATIONAL ERROR::YES	28.5373	LOSS OF CONTROL::NO	5.2939
LOSS OF CONTROL OF OTHER SHIP::YES	26.9525		
INCAPACITATED::REDUCED	26.1628		
OFF COURSE::YES	22.8674	observable risk factors	SoE (%)
VESSEL DAMAGE::YES	22.6119	INCAPACITATED::REDUCED	26,1628
OTHER SHIP ERROR::YES	18.8907	WIND FORCE::STORM	13,7905
WIND FORCE::STORM	13.7905	TIRED::YES	11,0279
STRESS LEVEL::HIGH	12.6865	FAMILARIZATION::NOT	9,269
TIRED::YES	11.0279	RADAR::NO	7,1763
FAMILIARIZATION::NOT	9.269	WEATHER::RAIN	6,0426
VISUAL DETECTION::POOR	7.3986	SAFETY CULTURE::POOR	4,1296
RADAR::NO	7.1763	SPECIAL CAUTION AREA::YES	2,3353
VISIBILITY::VERY POOR	6.5486	AIS::NO	1,358
WEATHER::RAIN	6.0426	SEA STATE::MODERATE	0,9797
BREAKDOWN::YES	4.8493	WIND DIRECTION::NE	0,2261
OTHER D.::MANY	4.6873		
SAFETY CULTURE::POOR	4.1296		
MAINTENANCE::NO	3.688		
SPECIAL CAUTION AREA::YES	2.3353		
AIS::NO	1.358		
SEA STATE::MODERATE	0.9797		
WIND DIRECTION::NE	0.2261		

According to the results obtained, the following recommendations for safety improvement during high traffic density can be distilled:

- Human related factors, their effects on situational awareness and the navigational error require special attention. In particular, the situational awareness and proper adjustment of navigational course are very significant in preventing unwanted hazard situations.
- The first line of intervention during high maritime traffic density is the prevention of the incapacitation and tiredness. Adequate cognitive and physical responses from the responsible seafarer importantly prevent the occurrence of accident through human error as the extremely strong influence factor identified by the developed approach.
- When in close encounter situation during high traffic density, both the knowledge on regulations for a give way and



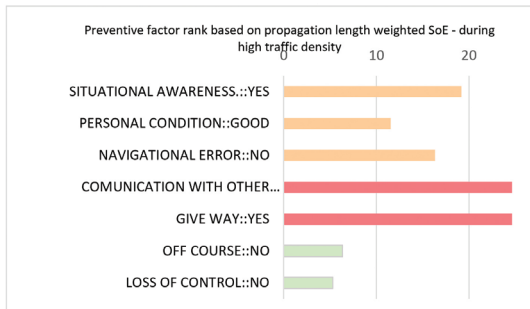


Figure 5.
Importance ranking of co-occurring collision influencing factors during high maritime traffic density based on propagation length weighted SoE.

the communication with other vessels, are critically important for the prevention of a collision within the analysed context.

5. CONCLUSION

Due to rare event characteristic of maritime accidents, the lack of data hinder the development of risk assessment models based on nowadays modern machine learning based models, and the Bayesian Network models has proven to be a good alternative solution that can embody sparse data, expertise, and experience in the maritime domain. However, when these probabilistic models of maritime accidents are put into work, the question of how to interpret the inference results and make it available to a wider group of interested experts without requiring the background knowledge in probabilistic methodology arises. The two aspects, modelling and interpretative reasoning, have been addressed in this paper. First, the Bayesian framework is exploited to develop a probabilistic causal model of maritime accident based on conceptual formulation of a causal network as a Bayesian Network classifier. The major strength of binary formulation of model outcomes lies in the possibility of introduction of weighting of evidence based on outcome hypothesis likelihood ratios. Along this line, the strength of evidence measure is derived and grading of results into influential categories is proposed. Thus, the strength of influence of the state space (influence factors) on the collision occurrence can be interpreted semantically without a background knowledge.

Though the complete framework is showcased for the modelling and the assessment of risks of collision for small vessels in navigation in the Adriatic, generalization can be made to other accident types in the maritime domain. The results of identified risks and preventive factors are obtained through the analysis of the conjunctive effect of factors and ultimately presented as the safety guarding recommendations in the scenario of navigation in high traffic density. Some recommendations obtained from the developed system are very intuitive to human reasoning, others are not as obvious, and become comprehensible and noticeable only when organized into a hierarchically structured influence diagram and after being quantitively evaluated. Therefore, the developed approach contributes to the focused

reasoning required for an intervention development under the given scenario. Similarly, any scenario can be imposed within the defined state space of the assessed accident.

It is important to mention the limitations that impact the results. The incomplete, insufficient, and scarce data is the universal problem for maritime risk assessment. It affects both the model development and the model validation. Consequently, it affects the inference results. The attention should be drawn to the uncertainty of network structure and the node parameters of the proposed model which represent the causal contributors to an accident. The network structure uncertainty should be investigated in future work. The maritime risk assessment obtained with the proposed weight of evidence approach and the recommendations distilled from the results should be viewed within the scope and limitations of the model. For a more in-depth overview on limitations of BNs in maritime accident domain, the interested reader is referred to (Hänninen, 2014).

Distances of the evidence nodes and target note play an important role in influence ranking for estimation of interventional priorities, but not in the classification of causative and preventive influences. According to (Yuan & Druzdzel, 2007), propagation length values are highly network dependant. Thus, an interesting direction of future work would be to investigate optimal values of the importance ranking based on propagation length, where not only risks are assessed, but interventional priorities are to be estimated as well. Also, in the future work, situational awareness should be researched in more detail as it is a very significant accident influencing factor that is affected by both human factors and technical equipment. More complex situational awareness subnetwork, that would accentuate the interplay of human factors and technical equipment, could yield a novel insight into its contributions to the occurrence of maritime accident, but also its potential for their prevention.

REFERENCES

Baker, C.C. et al., 2005. Accident Database Review of Human-Element Concerns: What Do The Results Mean For Classification? Human Factors In Ship Design, Safety & Operation. Available at: http://dx.doi.org/10.3940/rina.hf.2005.11.

Baksh, A.-A. et al., 2018. Marine transportation risk assessment using Bayesian Network: Application to Arctic waters. Ocean Engineering, 159, pp.422–436. Available at: http://dx.doi.org/10.1016/j.oceaneng.2018.04.024.

BayesFusion, L., 2020. Pittsburgh, USA. Available at: www.bayesfusion.com.

Bielza, C. & Larrañaga, P., 2014. Discrete Bayesian Network Classifiers. ACM Computing Surveys, 47(1), pp.1–43. Available at: http://dx.doi.org/10.1145/2576868.

Chan, H. & Darwiche, A., 2002. Reasoning about bayesian network classifiers. Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI2003), Acapulco, Mexico, Morgan Kaufmann, pp. 107-115.

Chauvin, C. & Lardjane, S., 2008. Decision making and strategies in an interaction situation: Collision avoidance at sea. Transportation Research Part F: Traffic Psychology and Behaviour, 11(4), pp.259–269. Available at: http://dx.doi.org/10.1016/j.trf.2008.01.001.

Chauvin, C. et al., 2013. Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. Accident Analysis & Prevention, 59, pp.26–37. Available at: http://dx.doi.org/10.1016/j.aap.2013.05.006.

Chen, P. et al., 2019. Probabilistic risk analysis for ship-ship collision: State-of-the-art. Safety Science, 117, pp.108–122. Available at: http://dx.doi.org/10.1016/j.ssci.2019.04.014.

Chen, P., Mou, J. & van Gelder, P.H.A.J.M., 2019. Integration of individual encounter information into causation probability modelling of ship collision accidents. Safety Science, 120, pp.636–651. Available at: http://dx.doi.org/10.1016/j.ssci.2019.08.008.

Chen, S.H. & Pollino, C.A., 2012. Good practice in Bayesian network modelling. Environmental Modelling & Software, 37, pp.134–145. Available at: http://dx.doi.org/10.1016/j.envsoft.2012.03.012.

Cockcroft, A. N. & Lameijer, J. N. F., 2003. Guide to the collision avoidance rules. 6th ed. Oxford: Butterworth-Heinemann.

Darwiche, A., 2009. Modeling and Reasoning with Bayesian Networks. Cambridge university press.

Friedman, N., Geiger, D. & Goldszmidt, M., 1997. Bayesian Network Classifiers. Machine Learning, 29(2/3), pp.131–163. Available at: http://dx.doi.org/10.1023/a:1007465528199.

Gaag, L. C. v. d., Bodlaender, H. L. & Feelders, A., 2004. Monotonicity in Bayesian networks. Proceedings of the 20th conference on Uncertainty in artificial intelligence, July 7-11, Banff, Canada, AUAI Press, pp. 569-576.

Goerlandt, F. & Montewka, J., 2015. Maritime transportation risk analysis: Review and analysis in light of some foundational issues. Reliability Engineering & System Safety, 138, pp.115–134. Available at: http://dx.doi.org/10.1016/j.ress.2015.01.025.

Gonzales, C., Torti, L. & Wuillemin, P.-H., 2017. aGrUM: A Graphical Universal Model Framework. Lecture Notes in Computer Science, pp.171–177. Available at: http://dx.doi.org/10.1007/978-3-319-60045-1_20.

Good, I. J., 1985. Weight of evidence: A brief survey. Bayesian statistics 2: Proceedings of the Second Valencia International Meeting, September 6-10, Amsterdam, pp. 249-269. Elsevier.

Hänninen, M. & Kujala, P., 2012. Influences of variables on ship collision probability in a Bayesian belief network model. Reliability Engineering & System Safety, 102, pp.27–40. Available at: http://dx.doi.org/10.1016/j.ress.2012.02.008.

Hänninen, M., 2014. Bayesian networks for maritime traffic accident prevention: Benefits and challenges. Accident Analysis & Prevention, 73, pp.305–312. Available at: http://dx.doi.org/10.1016/j.aap.2014.09.017.

Hassall, K.L. et al., 2019. Facilitating the elicitation of beliefs for use in Bayesian Belief modelling. Environmental Modelling & Software, 122, p.104539. Available at: http://dx.doi.org/10.1016/j.envsoft.2019.104539.

Huang, Y. et al., 2020. Ship collision avoidance methods: State-of-the-art. Safety Science, 121, pp.451–473. Available at: http://dx.doi.org/10.1016/j.ssci.2019.09.018.

Jeffreys, H., 1998. Theory of probability. 3rd reprint ed. Oxford: Oxford University Press.

Jensen, F. V. & Nielsen, T. D., 2007. Bayesian Networks and Decision Diagrams. Springer-Verlag New York.

Kass, R.E. & Raftery, A.E., 1995. Bayes Factors. Journal of the American Statistical Association, 90(430), pp.773–795. Available at: http://dx.doi.org/10.1080/01621459.1995.10476572.

Kjærulff, U. & Gaag, L. C. v. d., 2000. Making Sensitivity Analysis Computationally Efficient. Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI2000), June 30 - July 3, Stanford, USA, pp. 317-325.

Kontovas, C.A. & Psaraftis, H.N., 2009. Formal Safety Assessment: A Critical Review. Marine Technology and SNAME News, 46(01), pp.45–59. Available at: http://dx.doi.org/10.5957/mtsn.2009.46.1.45.

Lim, G.J. et al., 2018. Models and computational algorithms for maritime risk analysis: a review. Annals of Operations Research, 271(2), pp.765–786. Available at: http://dx.doi.org/10.1007/s10479-018-2768-4.

Madsen, A.L. & Jensen, F.V., 1999. Lazy propagation: A junction tree inference algorithm based on lazy evaluation. Artificial Intelligence, 113(1-2), pp.203–245. Available at: http://dx.doi.org/10.1016/s0004-3702(99)00062-4.

Marcot, B.G. & Penman, T.D., 2019. Advances in Bayesian network modelling: Integration of modelling technologies. Environmental Modelling & Software, 111, pp.386–393. Available at: http://dx.doi.org/10.1016/j.envsoft.2018.09.016.

Mazaheri, A., Montewka, J. & Kujala, P., 2016. Towards an evidence-based probabilistic risk model for ship-grounding accidents. Safety Science, 86, pp.195–210. Available at: http://dx.doi.org/10.1016/j.ssci.2016.03.002.

Ministry of Sea, T. a. I. & Tourism, M. o., 2008. Nautical tourism development strategy of the Republic of Croatia 2009–2019.[Online] Available at: mmpi.gov.hr/UserDocsImages/arhiva/Strategija%20razvoja%20nautickog%20turizma%20ENGL%201.pdf.

Mukhopadhyay, N., 2020. Probability and Statistical Inference. Available at: http://dx.doi.org/10.1201/9780429258336.

Osteyee, D. B. & Good, I. J., 1974. Information, weight of evidence, the singularity between probability measures and signal detection (Lecture Notes in Mathematics), Springer-Verlag Berlin Heidelberg.

Øvergård, K. I., 2015. Human Error: Causality and the Confusion of Normative and Descriptive Accounts of Human Performance. In K. I. Fostervold, S. Å. K. Johnsen, L. Rydstedt and R. G. Watten (Eds.) Creating Sustainable Work Environments, (pp. C5-6, C5-10). Lysaker, Norway: Norwegian society for ergonomics and human factors. Available at: https://openarchive.usn.no/usn-xmlui/handle/11250/2486731.

Øvergård, K.I., Tannum, M.S. & Haavardtun, P., 2020. Observation and assessment of crossing situations between pleasure craft and a small passenger ferry. WMU Journal of Maritime Affairs, 19(3), pp.337–355. Available at: http://dx.doi.org/10.1007/s13437-020-00211-1.

Ozturk, U. & Cicek, K., 2019. Individual collision risk assessment in ship navigation: A systematic literature review. Ocean Engineering, 180, pp.130–143. Available at: http://dx.doi.org/10.1016/j.oceaneng.2019.03.042.

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. San Francisco, CA,United States: Morgan Kaufmann.

Pearl, J., 2000. Causality: Models, Reasoning, and Inference. 2nd Ed. ed. Cambridge, UK: Cambridge University Press.

Pianosi, F. et al., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software, 79, pp.214–232. Available at: http://dx.doi.org/10.1016/j.envsoft.2016.02.008.

Plajner, M. & Vomlel, J., 2017. Monotonicity in Bayesian Networks for Computerized Adaptive Testing. Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp.125–134. Available at: http://dx.doi.org/10.1007/978-3-319-61581-3_12.

Rausand, M. & Høyland, A., 2003. System Reliability Theory: Models, Statistical Methods, and Applications. 2nd Ed., Wiley.

Sigurdsson, J.H., Walls, L.A. & Quigley, J.L., 2001. Bayesian belief nets for managing expert judgement and modelling reliability. Quality and Reliability Engineering International, 17(3), pp.181–190. Available at: http://dx.doi.org/10.1002/qre.410.

Sotiralis, P. et al., 2016. Incorporation of human factors into ship collision risk models focusing on human centred design aspects. Reliability Engineering & System Safety, 156, pp.210–227. Available at: http://dx.doi.org/10.1016/j.ress.2016.08.007.

Stornes, P., 2015. Risk influencing factors in maritime accidents : An exploratory statistical analysis of the Norwegian Maritime Authority incident database. [Online] Available at: https://samforsk.brage.unit.no/samforsk-xmlui/handle/11250/2360509.

Trucco, P. et al., 2008. A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering & System Safety, 93(6), pp.845–856. Available at: http://dx.doi.org/10.1016/j.ress.2007.03.035.

Yuan, C. & Druzdzel, M.J., 2006. Importance sampling algorithms for Bayesian networks: Principles and performance. Mathematical and Computer Modelling, 43(9-10), pp.1189–1207. Available at: http://dx.doi.org/10.1016/j.mcm.2005.05.020.

Zhang, D. et al., 2013. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. Reliability Engineering & System Safety, 118, pp.93–105. Available at: http://dx.doi.org/10.1016/j.ress.2013.04.006.

Zhang, G. & Thai, V.V., 2016. Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. Safety Science, 87, pp.53–62. Available at: http://dx.doi.org/10.1016/j.ssci.2016.03.019.

Zhang, J. et al., 2018. Quantitative assessment of collision risk influence factors in the Tianjin port. Safety Science, 110, pp.363–371. Available at: http://dx.doi.org/10.1016/j.ssci.2018.05.002.

APPENDIX

Variables	Description	State	Data source
Daytime	Statistical number of maritime accidents during the day or night. In the case of passenger cruise ships, the risk of a maritime accident is higher during the day, since the observed ships are more active during the day.	day/night	MMPI [1]
Season	Indicates whether the marine casualty occurred in summer or winter. The summer period is period from April 1 to October 31.	summer/winter	MMPI
Day	Refers to the relationship between the probability of a maritime accident and the day of the week.	7	MMPI
Weather	Describes a meteorological condition during a maritime accident	4	MMPI
Sea state	Refers to the state of the sea at the time of the maritime accident	5	MMPI
Wind force	Refers to the wind force at the time of the maritime accident	6	MMPI
Wind direction	Refers to the wind direction at the time of the maritime accident	3	MMPI
Visibility	Refers to the visibility at the time of the maritime accident	5	MMPI
Traffic distribution	The ratio of the total number of vessels and the area of the observed area, where the term vessel includes ships, yachts and boats	low/high	MMPI1 [2]
Visual detection	Refers to the visibility from the bridge which is conditioned by the design and arrangement of the windows on the bridge, the wipers on the windows, the salt on the windows, etc. The variable also refers to the visibility conditioned by the intensity of the sunlight.	good/bad	expert
Tired	Refers to the condition of the person operating the ship.	yes/no	DNV1 [3],expert
Familiarization	Refers to the experience of a person operating a ship sailing a certain area.	yes/no	DNV [4],expert
Stress level	Refers to how much stress the person operating the ship is exposed to.	yes/no	DNV,expert
Incapacitated	Refers to the mental ability of the person operating the ship. Disability can occur due to effects of alcohol, illness, drug abuse, or some medications.	yes/no	DNV,expert
Other distractions	Refers to the exposure of the person operating the ship to other distractions, such as mobile devices, the presence of other people on the bridge, problematic situations on the ship that may distract him from navigation tasks.	yes/no	DNV, expert
Situational awareness	Refers to a person's ability to construct a mental model based on the present status, and make projections into the future environment, both onboard and around	yes/no	DNV, expert
Personal condition	Refers to the mental and physical condition of the person operating the ship.	good/bad	expert

1. *MMPI, data obtained per request from the databases of The Ministry of Maritime Affairs, Transport and Infrastructure of the Republic of Croatia, https://mmpi.gov.hr/*
2. *MMPI1, Available at.: https://mmpi.gov.hr/UserDocsImages/arhiva/MMPI%20-%20South%20Adriatic%20v.3.1%2022-12_14.pdf*
3. *DNV1, Det Norske Veritas, Formal safety assessment of Electronic Chart Display and Information System (ECDIS). Technical report. DNV; 2006.*
4. *DNV, Det Norske Varitas, Formal safety assessment – large passenger ships, technical report, DNV, 2003*

Communication with other vessel	Refers to successful communication with another ship in the event of collision courses or during navigation through narrow channels	good/bad	expert
Safety culture	- possible lack of crew and human resources on board, poor financial situation of the company that owns the ship, frequent changes of crew due to dissatisfaction with working conditions - working hours, i.e. the possibility for seafarers to work in shifts that can result in increased fatigue and reduced quality of health. - non-compliance with safety protocols, poor awareness of the crew about maritime safety.	high/standard/ bad	DNV,expert
Give way	Refers to taking proper actions to avoid collisions in accordance with the "Pravilnik o sigurnosti pomorske plovidbe u unutarnjim morskim vodama i teritorijalnom moru Republike Hrvatske te načinu i uvjetima obavljanja nadzora i upravljanja pomorskim prometom.",(Croatian Maritime Law,2013)[5]	yes/no	expert
Maintenance	Refers to the maintenance of technical systems, ship hull and ship systems in general.	good/bad	expert
AIS	Refers to correct use of the automatic identification system on board	yes/no	expert
Radar	Refers to the correct use of the radar on a ship	yes/no	expert
Vessel damage	Refers to all damage to the ship regardless of cause and effect.	yes/no	expert
Special caution area	Refers to areas of special danger specific to the observation area. Areas of special caution defined according to the Maritime-Navigation Study of the Split, Ploče and Dubrovnik navigable areas are: Splitska vrata, Drvenički kanal, Viški kanal, Šoltanski kanal, Pakleni otoci and Pakleni kanal, Prilaz Gradskoj luci, areas of seaplane navigation, areas of the outer edges of the islands.	yes/no	expert
Loss of control	Refers to the loss of control of a ship due to a technical malfunction or due to human error during which nothing can stop the ship from moving towards danger.	yes/no	DNV, expert
Breakdown	Refers to a technical failure on board, regardless of the cause and effect of the failure.	yes/no	DNV, expert
Human error	Refers to intolerant activity or deviation from normal behaviour whose boundaries are defined by the system.	yes/no	DNV, expert
Navigational error	Refers to errors in navigating the sea, i.e. in determining the position of the ship, control and supervision of its movement.	yes/no	DNV, expert
Off course	Refers to a group of causes that lead to the ship being unable to navigate at the planned course.	yes/no	DNV, expert
Loss of control of other ship	Refers to the loss of control of another ship due to a technical or human error, or the influence of external factors.	yes/no	expert
Other ship error	Refers to human errors on another ship (navigation errors, handling errors, etc.)	yes/no	expert

5. *Croatian Maritime Law, 2013., Available at: https://narodne-novine.nn.hr/clanci/sluzbeni/2013_06_79_1640.html*