Risk Analysis, Vol. 26, No. 3, 2006 DOI: 10.1111/j.1539-6924.2006.00775.x

Enabling a Powerful Marine and Offshore Decision-Support Solution Through Bayesian Network Technique

A. G. Eleye-Datubo,[1] A. Wall,[1] A. Saajedi,[1] and J. Wang[1*]

A powerful practical solution is by far the most desired output when making decisions under the realm of uncertainty on any safety-critical marine or offshore units and their systems. With data and information typically being obtained incrementally, adopting Bayesian network (BN) is shown to realistically deal with the random uncertainties while at the same time making risk assessments easier to build and to check. A well-matched methodology is proposed to formalize the reasoning in which the focal mechanism of inference processing relies on the sound Bayes's rule/theorem that permits the logic. Expanding one or more influencing nodal parameters with decision and utility node(s) also yields an influence diagram (ID). BN and ID feasibility is shown in a marine evacuation scenario and that of authorized vessels to floating, production, storage, and offloading collision, developed via a commercial computer tool. Sensitivity analysis and validation of the produced results are also presented.

1. INTRODUCTION

If all the information that could be known about a maritime hazardous event/situation were obtainable for its risk assessment, then the results of such studies that are accurately carried out would not be subject to uncertainty. Instead, data and information are typically obtained incrementally. Thus, the inherent uncertainty can be due to imperfect understanding of the domain, incomplete knowledge of the state of the domain at the time where a given task is to be performed, randomness in the mechanisms governing the behavior of the domain, or a combination of these. It is necessary then to model the assessment domain such that the probabilistic measure of each event becomes more reliable in light of the new information being received. In view of this, the domain that is represented can be put out in an intuitive visual format as a Bayesian network (BN) model. The BN reasoning system can be viewed as the generalization of prepositional logic and resolution theorem proving that incorporates the treatment of uncertainty for the structure of the complex argument. Probability and Bayes's theory ensure that inferences based on the network are sound.

As essential in a risk-based marine community, reasoning with incomplete knowledge is one of the fundamental features of human intelligence. Competent expert and engineering judgment (to compensate for any lack of mature data) incorporated in a BN can aid in providing its solid knowledge base. The generic nature of this technique means that it can be developed further and applied widely in marine and offshore applications. With this philosophy in a logical framework, adopting BN to formalize reasoning about system dependability will make assessments easier to build, check, and certainly update.

The analogy of BN models can be further expanded/transformed to output influence diagrams (IDs) that are highly intuitive in the decision-making process. Such diagrams aid the visibility of a large number of interacting issues and their effects on the

[1] Marine, Offshore & Transport Research Group, School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK.

[*] Address correspondence to Jin Wang, Liverpool John Moores University; tel: +44 151 231 2445; fax: +44 151 231 2453; j.wang@ljmu.ac.uk.

decision. They can also offer the benefit of a robust practical solution that is required for achieved safety at an affordable cost. Hence, the final scheme of the BN can give a model in which reasoning is justified, while it enables a powerful marine decision-support solution that is easy to use, flexible, and appropriate for the assessment task.

2. LITERATURE BRIEFING ON BN

Until 20 years ago, the issue of ordering possible beliefs, both for belief revision and for action selection, was seen as increasingly important and problematic, and at the same time, dramatic new developments in computational probability and decision theory directly addressed perceived shortcomings. The key development (Pearl, 1988) was the discovery that a relationship could be established between a well-defined notion of conditional independence in probability theory and the absence of arcs in a directed acyclic graph (DAG). This relationship made it possible to express much of the structural information in a domain independent of the detailed numeric information, in a way that both simplifies knowledge acquisition and reduces the computational complexity of reasoning. The resulting graphic models have come to be known as BNs.

BNs are at the cutting edge of expert systems research and development. Unlike the traditional rule-based approach to expert systems, they are able to replicate the essential features of plausible reasoning (reasoning under conditions of uncertainty) and combine the advantages of an intuitive visual representation with a consistent, efficient, and mathematical basis in Bayesian probability. Critically, they are capable of retracting belief in a particular case when the basis of that belief is explained away by new evidence. Because of the development of propagation algorithms (Lauritzen & Spiegelhalter, 1988; Pearl, 1988; Russell & Norvig, 2003), followed by availability of easy-to-use commercial software and growing number of creative applications (Jensen, 1993; SERENE Consortium, 1999), BN has caught the sudden interest of research in different research fields since the early 1990s. Perhaps the greatest testament to the usefulness of Bayesian problem-solving techniques is the wealth of practical applications that have been developed since then in areas of intelligent decision, safety assessment, information filtering, autonomous vehicle navigation, weapons scheduling, medical diagnosis, pattern recognition, and computer network diagnosis (Heckerman *et al.*, 1995). Since most real-life problems involve inherently uncertain relationships, BN is a technology with huge potential for application across many domains.

IDs, which further extend the notion of BNs by including decision nodes and utility nodes, have been used in human reliability assessment (Humphries, 1995) and decision making on explosion protection offshore (Bolsover & Wheeler, 1999). A good reference work for the computational method underlying the implementation of them in *Hugin* is described in (Jensen *et al.*, 1994). The *Hugin* software (Jensen, 1993) enables a powerful risk assessment solution that is easy to use, flexible, and appropriate for use on marine and offshore applications. Other renowned program packages for BN building and influencing include *MSBNx* (Kadie *et al.*, 2001), created at Microsoft Research, and *Netica* (Netica, 2002), the commercial program developed by Norsys Software Corp.

3. SEMANTICS OF A BN

Fundamental to the idea of BNs is the concept of modularity, whereby a complex system is built by combining simpler parts of components that are related in a causal manner. A BN provides factorized representation of a probability model that explicitly captures much of the structure typical in human-engineered models. More generally, a BN is a DAG that encodes a conditional probability distribution (CPD) at its nodes on the basis of arcs received. Therefore, by definition:

$$\text{BN} = \text{DAG encoded with CPD.}$$

The graphical structure of a BN (i.e., the DAG) depicts a qualitative illustration of the interactions among the set of random (i.e., chance) variables, such as hazardous events, that it models. Numerically, a BN represents the joint probability distribution (JPD) among the modeled variables. This distribution is described efficiently, exploring probabilistic independencies among the modeled variables. Each node is described by a probability distribution (PD) conditional on its direct predecessors that has its values entered into a conditional probability table (CPT), i.e., a *matrix of conditional probabilities*, associated with the node. The encoded nodes with no predecessors are described by *prior* PDs. Those with predecessors are described by *posterior* PDs.

4. BAYESIAN INFERENCE MECHANISM

Bayesian inference is a process by which observations of a real-world situation are used to update the random uncertainty about one or more variables characterizing aspects of that situation. It relies on the use of *Bayes's rule/theorem* (Bayes, 1763) as its rule of inference, defining the manner in which uncertainties ought to change in light of newly made observations. This subjective probability theory is only part of the Bayesian inference mechanism. Together with the applicable results of such probability concepts as the product and sum rules, the concept of conditional independence (Pearl, 1988), dependency separated or d-separated (Pearl, 1988), the techniques of marginalization (Vellido & Lisboa, 2001), and the pattern of inference (Wellman & Henrion, 1993; Lauritzen & Spiegelhalter, 1988; Pearl, 1988), it provides the basic tool for both Bayesian belief updating and for treating probability as logic. In order to apply these tools, the prior probabilities and the likelihood probabilities must be obtained.

4.1. Bayes's Theorem/Rule

In order to make probability statements about the model parameters, the analysis must begin with providing initial or *prior probability* estimates for specific outcomes or events of interest. Then from sources such as a special report, a database, a case study, etc., some additional information (i.e., data or evidence) about the event, or an entirely new event(s), is obtained. In light of this new information providing new data belief, it is desirable to improve the state of knowledge, and thus the prior probability values are updated by calculating revised probabilities, referred to as the *posterior probabilities* (these probabilities provide the basis for action). *Bayes's theorem* provides a means for making these probability calculations. Essentially, it is a relationship between conditional and marginal probabilities, and is given for two events, A and B, by Equation (1).

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}. \tag{1}$$

Each term in Bayes's theorem has a conventional name. The term $P(A)$ is called the *prior probability of* A. It is "prior" in the sense that it precedes any information about B and this is what causes all the arguments. $P(A)$ is also the *marginal (total) probability of* A. The term $P(A \mid B)$ is called the *posterior probability of A, given B*. It is "posterior" in the sense that it is derived from or entailed by the specified value of B. The term $P(B \mid A)$, for a specific value of B, is called the *likelihood function for A, given B* and can also be written as $L(A \mid B)$. The term $P(B)$ is the *prior* or *marginal (total) probability of B,* but also one that provides *evidence* of interest for the probability update of A. Its inverse is usually regarded as a *normalizing constant*, α. With this terminology, the theorem may be paraphrased as

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} \rightarrow P(A \mid B) = \alpha L(A \mid B) P(A). \tag{2}$$

Generally, for an event B with states $\{b_1, \ldots, b_m\}$, the posterior probability on the event A can be computed from the Bayes's rule as

$$P(A \mid b_1, \ldots, b_m) = \frac{P(b_1, \ldots, b_m \mid A)P(A)}{P(b_1, \ldots, b_m)}. \tag{3}$$

The process of Bayes's theorem is repeated every time new or additional information becomes available, so that as Lindley (1970) puts it, "today's posterior probability is tomorrow's prior." Thus, as the number of pieces of evidence increases, the dependence of the posterior on the original estimated prior decreases.

4.2. The Likelihood Function

The *likelihood principle* (Fisher, 1922; Edwards, 1992) states that *all* the relevant information in the model is contained in the likelihood function (which is of fundamental importance in the theory of Bayesian inference). "Likelihood" is a solitary term used to represent such a function and is one of several informal synonyms for "probability"; so sometimes, $P(B \mid A)$ is called the *likelihood of A, given B*, and is denoted by $L(A \mid B)$. The reason for this is that if, for example, $a_1, \ldots, a_n$ are possible states of event A with an effect on the event B in which b is known, then $P(b \mid a_i)$ is a measure of how likely it is that a_i is the cause. Moreover, this is a simple, compelling concept that has a host of good statistical properties and can be derived from the reasoning logic as well as by expert judgment.

5. STRUCTURAL EFFECTS ON THE INFERENCE PROCESSING

One of the best features of BNs is that one can incorporate *new node*(s) as the data become available. Thus, it follows that one "effect" can be a "cause" of a new/another node and a "cause" can also be the "effect" of a new/another node. Owing to this additional capability of a BN model, it can constitute a description of the probabilistic relationships among the system's variables that amount to a factorization of the joint distribution of all variables into a series of marginal and conditional distributions. Evidence propagation may take place via a message-posting scheme.

5.1. Joint Probability Distribution

A probabilistic model may consist of a set of variables $X = \{X_1, X_2, \ldots, X_n\}$, which exploits *conditional independence* to represent the JPD over X having the product form (Pearl, 1988):

$$\begin{aligned} P(x_1, \ldots, x_n) &= P(x_1 \mid \mathrm{parent}(X_1)) \\ &\quad \times P(x_2 \mid \mathrm{parent}(X_2)) \ldots P(x_n \mid \mathrm{parent}(X_n)) \\ &= \prod_{i=1}^{n} P(x_i \mid \mathrm{parents}(X_i)). \end{aligned} \tag{4}$$

$P(x_1, x_2, \ldots, x_n)$ gives the JPD and, like the CPD, it is a table of values where one entry is made for each possible combination of values that its variables can jointly take. The JPD for a problem captures the probability information of every possible combination of a set of variables, and their states. Once a JPD has been defined for a problem, then it is possible, using it along with the *axioms of probability*, to answer any probabilistic query regarding any of the variables. This includes their value given additional evidence, that is, their posterior probabilities, although the space, and consequently, time complexity required in representing and manipulating the JPD is exponential in the number of variables considered (D'Ambrosio, 1999). For example, the JPD required to represent a system with 20 binary values would have 2^{20} (1,048,576) values. This causes a problem in the elicitation, storage, and manipulation of these values, thus making the use of JPDs unfeasible for any practical use. Fortunately, when modeling most real systems, advantage is taken of any inherent structure the system has by modeling the system as a graph (D'Ambrosio, 1999).

In the general case, a JPD over a set of variables, $X = \{X_1, X_2, \ldots, X_n\}$, can be defined recursively using the product rule (Equation (5)):

$$\begin{aligned} &P(X_1, X_2, \ldots, X_n) \\ &= P(X_1 \mid X_2, \ldots, X_n) P(X_2, \ldots, X_n) \\ &= P(X_1 \mid X_2, \ldots, X_n) P(X_2 \mid X_3, \ldots, X_n) P(X_3, \ldots, X_n) \\ &= P(X_1 \mid X_2, \ldots, X_n) \\ &\quad \times P(X_2 \mid X_3, \ldots, X_n) \cdots P(X_{n-1} \mid X_n) P(X_n). \end{aligned} \tag{5}$$

This factorization property of JPDs is referred to as the *chain rule* of probabilities and is one that allows any ordering of variables in the factorization. Such a rule is especially significant for BNs, because it provides a means of calculating the full JPD from conditional probabilities, which is what a BN stores. For example, the JPD for three events, A, B, and C, can be expressed more compactly as:

$$\begin{aligned} P(A \mid B, C) P(B, C) &= P(A, B, C) \\ &= P(B \mid A, C) P(A, C). \end{aligned} \tag{6}$$

Then, in applying Equation (5), Bayes's theorem specifies the probability of an event A, given the condition that an event B and an event C both occur ($B \rightarrow A \leftarrow C$) as:

$$P(A \mid B, C) = \frac{P(B \mid A, C) P(A \mid C)}{P(B \mid C)}. \tag{7}$$

5.2. Belief Probability Update

Evidence is new information about a random variable that causes a change about its PD. Newly available evidence is brought about when a particular state of an event happens. The effect of such new evidence will certainly propagate throughout the network and thereby cause the posterior probabilities of other events to iteratively be recalculated. This is achievable by *message posting* along the edges (Pearl, 1988). Therefore, introducing the notion of evidence is imperative in the reasoning with BN. Nonetheless, it is worth noting that the real power and generalization of BN is that entered evidence propagates in both directions, even though the graph is directed.

Suppose there is an interest in a given event C (referred to as the *query variable*) having a joint probability $P(c)$, over C. Before any evidence becomes available, the propagation process consists of calculating the marginal probabilities $P(C_i = c_i)$, or simple $P(c_i)$, for each C_i.

Now, suppose some evidence has become available to the event C. In this situation, the propagation process consists of calculating the conditional probabilities $P(C_i = c_i \mid \varepsilon = e)$, or simple $P(c_i \mid e)$, where ε is a set of *evidential nodes* with known values $\varepsilon = e$.

The newly available evidence, ε, can be decomposed into two subsets:

- ε_i^+, the subset of ε that can be accessed from C_i though its parents (top-down), i.e., propagates in the direction of the arcs.
- ε_i^-, the subset of ε that can be accessed from C_i though its children (bottom-up), i.e., propagates against the direction of the arcs.

For the probability of $C_i = c_i$, given that $e = e_i^+$ for a parent and $e = e_i{}^-$ for a child:

$$P(c_i \mid e) = P\left(ci \mid e_i^-, e_i^+\right) = \frac{P\left(e_i^- \mid c_i, e_i^+\right) P\left(c_i \mid e_i^+\right)}{P\left(e_i^- \mid e_i^+\right)}. \tag{8}$$

Since C_i d-separates ε_i^- from ε_i^+ (i.e., $\varepsilon_i^- || \varepsilon_i^+$, where $||$ stands for d-separation), conditional independence can be used to simplify the first term in the numerator and then $1/P(e_i^- \mid e_i^+)$ can be treated as a normalizing constant, α, so that:

$$P(c_i \mid e) = \alpha P\left(e_i^- \mid c_i\right) P\left(c_i \mid e_i^+\right). \tag{9}$$

According to the Bayes's theorem conventional interpretation (Equation (2)), posterior is prior scaled by likelihood and normalized by evidence (so $\Sigma(\text{posteriors}) = 1$), thus Equation (9) can be rewritten as

$$P(c_i \mid e) = \alpha \lambda_i(c_i) \pi_i(c_i), \tag{10}$$

where

$\lambda_i(c_i)$ represents $P(e_i^- \mid c_i)$, a message passed onto c_i as *likelihood evidence*; and

$\pi_i(c_i)$ represents $P(c_i \mid e_i^+)$, a message passed onto c_i as *prior evidence*.

To compute the functions $\lambda_i(c_i)$ and $\pi_i(c_i)$, suppose a typical node C_i has parents $B = \{B_1, \ldots, B_m\}$ and children $A = \{A_1, \ldots, A_n\}$ (see Fig. 1).

The evidence ε_i^+ can be partitioned into m disjoint components, one for each parent of C_i:

$$\varepsilon_i^+ = \left\{\varepsilon_{B_1 C_i}^+, \ldots, \varepsilon_{B_m C_i}^+\right\}, \tag{11}$$

where the evidence $\varepsilon_{B_j C_i}^+$ is the subset of ε_i^+ contained in the B_j-side of the link $B_j \to C_i$.

Similarly, the evidence ε_i^- can be partitioned into n disjoint components, that is:

$$\varepsilon_i^- \left\{\varepsilon_{A_1 C_i}^-, \ldots, \varepsilon_{A_n C_i}^-\right\}, \tag{12}$$

where the evidence $\varepsilon_{B_j C_i}^-$ is the subset of ε_i^- contained in the A_j-side of the link $A_j \leftarrow C_i$.

Then, given an instantiation of $b = \{b_1, \ldots, b_m\}$ of the parents of C_i, $\pi_i(c_i)$ can be computed (i.e., *top-down propagation*) via a recursive solution (Pearl, 1986; Castillo *et al*., 1997). Likewise, given an instantiation of $a = \{a_1, \ldots, a_n\}$ of the children of C_i, $\lambda_i(c_i)$ can be computed (i.e., *bottom-up propagation*).

The CPTs of the events *never* change by entering new evidence; only the new-fangled/belief probability in each of its possible states is determined by the belief probability in the states of the nodes to which it is directly connected. The algorithm simultaneously updates belief for all the nodes, causing them to become posterior probabilities, until the network reaches equilibrium. In other words, the JPD of the variables changes each time new information is learnt about the observable variables. Such calculations for the propagation of probabilities in a BN are usually

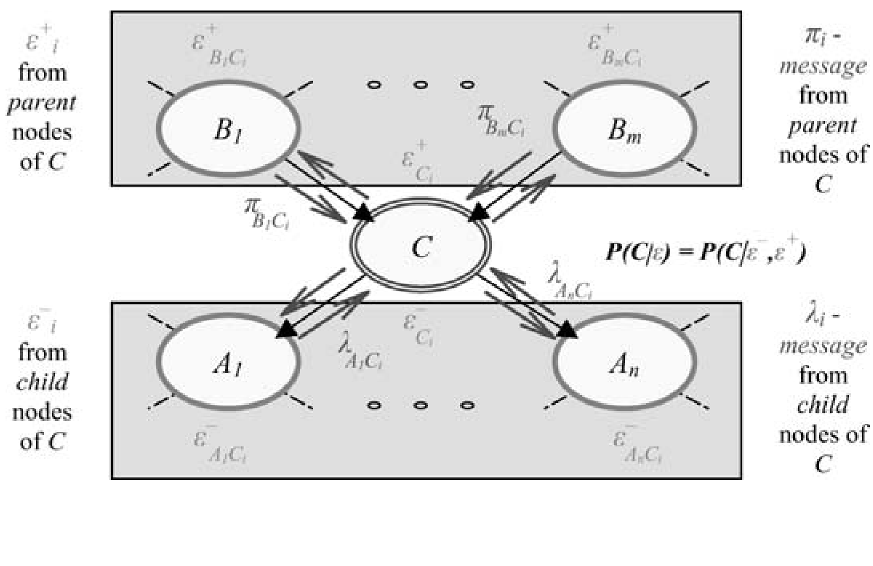


Fig. 1. Evidence propagation via message posting.

tedious (Jensen *et al.*, 1990). Therefore, *Hugin* is used as the robust BN programming environment for modeling and calculations (Jensen, 1993). This software tool allows for interactive creation of the network, maintenance of knowledge bases, and incorporates new, efficient algorithms to support the execution of Bayesian probability calculations, thus making a complete probabilistic model.

6. INFLUENCE DIAGRAM

An ID was originally a compact representation of a decision tree for a symmetric decision scenario: one is faced with a specific sequence of decisions, and between each decision one observes a specific set of variables. Nowadays, an ID is a BN expanded with *utility functions* and with *variables representing decisions*, in order to provide decision-making capabilities within the model. The utilities and decisions are both represented using nodes of distinguishing shapes in contrast to that of BN variables. In fact, the subset of an ID that consists of only *chance nodes* is a BN. Therefore, by definition:

$$\text{ID} = \text{BN expanded with decisions and utilities functions.}$$

An ID that uses only these elements is a simple but powerful communication tool, and one that can also be used to perform a quantified assessment of the decision problem. While the ID is very useful in showing the qualitative structure of the decision problem (Gámez *et al.*, 2004) for the domain, the network must also remain acyclic, and there must exist a directed path that contains all decision nodes (usually drawn as rectangles or squares) in the network. Decision-makers are interested in making the best possible decisions (i.e., the preferences) for an application, and therefore utilities are associated with the state configurations of the network. Each utility node (normally drawn as diamond-shaped or hexagons) has a utility function that associates a utility to each configuration of states of its parents (utility nodes do not have children).

Making decisions influences the probabilities of the configurations of the network and as such, the algorithms for probability updating can be modified to solving IDs. Its evaluation is done by setting the value of the decision node to a particular choice of action (i.e., best risk control option (RCO)), and treating the node just as a nature node with a known value that can further influence the values of other nodes. The action's utility is calculated first by calculating the conditional probabilities for the parents of the utility node using standard inference algorithm, and then feeding the results to the utility function. One can compute the expected utility (EU) of each decision alternative (the global utility function is the sum of all the local utility functions). The alternative with the highest EU is chosen; this is known as the maximum expected utility (MEU) principle.

6.1. Expected Utility

In order to assess the decision alternatives in D, a utility table $U(D, S)$ is needed to yield the utility for each configuration of decision alternative and outcome state for the determining variable. The EU of a given decision alternative d is calculated by

$$\text{EU}(d) = \sum_{S} P(S \mid d) U(d, S), \tag{13}$$

where $U(d, S)$ are the entries of the utility table in the value node U. The conditional probability $P(S \mid d)$ is computed from CPT of the determining variable having outcome states, $s \in S$, given that the decision alternative d is fired.

6.2. Maximum Expected Utility

There is the presumption from *utility theory* (Von Neumann & Morgenstern, 1964), and from *decision theory* as well (North, 1968; French, 1988), that humankind is *rational* when inferring subjective value (or utility) from choices (or preferences). This implies that decisionmakers maximize their utility wherever possible. Based on this, two principles are then used to determine the existence of the utility function:

- *Utility principle*: If a decisionmaker obeys the axioms of utility, then there exists a real-valued function, U, that operates on states such that $U(X) > U(Y)$ if and only if X is preferred to Y, and $U(X) = U(Y)$ if and only if there is no preference between X and Y.
- *MEU principle*: This implies that a rational decisionmaker should choose an action that maximizes EU of outcome states. Thus, given that $d_1, d_2, \ldots, d_k$ are the mutually exclusive decision alternatives of D, the decision alternative d that gives MEU is:

$$\text{MEU}(d) = \max_{d}\{\text{EU}(d_1), \text{EU}(d_2), \ldots, \text{EU}(d_k)\}. \tag{14}$$

Utility theory can be used in both *decision making under risk* (where the probabilities are explicitly given) and in *decision making under uncertainty* (where the probabilities are not explicitly given). The theory can be expanded to application for safety-based marine and offshore decisions through *cost-benefit evaluation*, whereby utmost considerations, for cost-effectiveness, are given to both *cost* and *safety* (i.e., *risk reduction*). In such a case, evaluation of RCOs according to values of implied cost of averting a fatality (ICAF), rather than the utility figures of an outcome state, may enable initial comparing and ranking of these options. The more attractive options for realization would be those with the lower ICAFs. The ability to map preferences (e.g., RCOs) into a single numerical value for ranking follows from the axioms of utility.

7. PROPOSED BN METHODOLOGY

A BN reasoning process has been developed to provide a natural framework for maritime risk assessment and decision support. A flow chart of the approach is shown in Fig. 2, and this format ensures that the BN analysis is conducted in a disciplined, well-managed, and consistent manner that promotes the delivery of quality maritime decision-making results. The depth or extent of application of the methodology should be commensurate with the nature and significance of the problem. Nonetheless, the entire methodology consists of nine key steps that have been encapsulated within the following three modules:

- Module 1: Visual BN Modeling (i.e., Steps 1 and 2).
- Module 2: Inference Algorithm of Bayesian Analysis (i.e., Steps 3 to 7).
- Module 3: Reasoning Evaluation via an ID (i.e., Steps 8 and 9).

In building a BN model, one can first focus on specifying the qualitative structure of the domain (Module 1) and then on quantifying the influences. When finished, one is guaranteed to have a complete specification of the PDs. Then, following evidence propagation (Module 2), an intuitive evaluation for decision making is enabled through added nodes of decisions and utilities (Module 3). *Hugin* is used as the robust BN programming environment for the risk modeling and its probability calculations. Explanations for each of the steps in these underlying modules are given as follows:

Step 1—Setting of Domain for Accident Category Information: Very important to the BN process is available information and failure data collected from every possible source, especially those from regulatory practice, databases and networks, tests, experiments, physical models, simulations, and analytical models. Expert judgment is utilized throughout the understanding of the domain and also in assigning valuable figures where data are not available. As observed data becomes available, they can be used to update, refine, or replace the estimates provided by subject matter experts. In this sense then, whenever there are uncertainties, e.g., in respect of data or expert judgment, the significance of these uncertainties and limitations will be identified, so as to assess the degree of reliance that should be placed on the available data.

Step 2—Creation of Nodes and Establishment of Probabilistic Relations: For the first step in constructing the BN, the development of the graphical representation, indicating the relevant variables (nodes) and dependencies (arcs), is important, not only because it determines the level of detail to be used in the subsequent functional model building, but also because it provides a straightforward means of analyzing and communicating causal assumptions that are not easily expressed using standard mathematical notation (Pearl, 2000).

In general, the problem under consideration is characterized by a number of functions or parameters (i.e., the relevant variables). These relate to, for example, a cause event, A, or an effect event, B, and can be mapped as labeled nodes into the network pane. Identified influence relationships between nodes are established such that an arc connection is placed between an influencing (parent) node and an influenced (child) node. The terminating arrowhead of the arcs is then set to point at the child nodes.

Step 3—Formulation of CPTs and Prior Probabilities: The inference consists of computing the conditional probabilities with the BN; thus, the next step will be to specify the states and to input values for a CPT (i.e., the conditional probability matrix) of each node. In other words, evidence can be entered to the network by manually setting probabilities in the network. The result of the associated tables

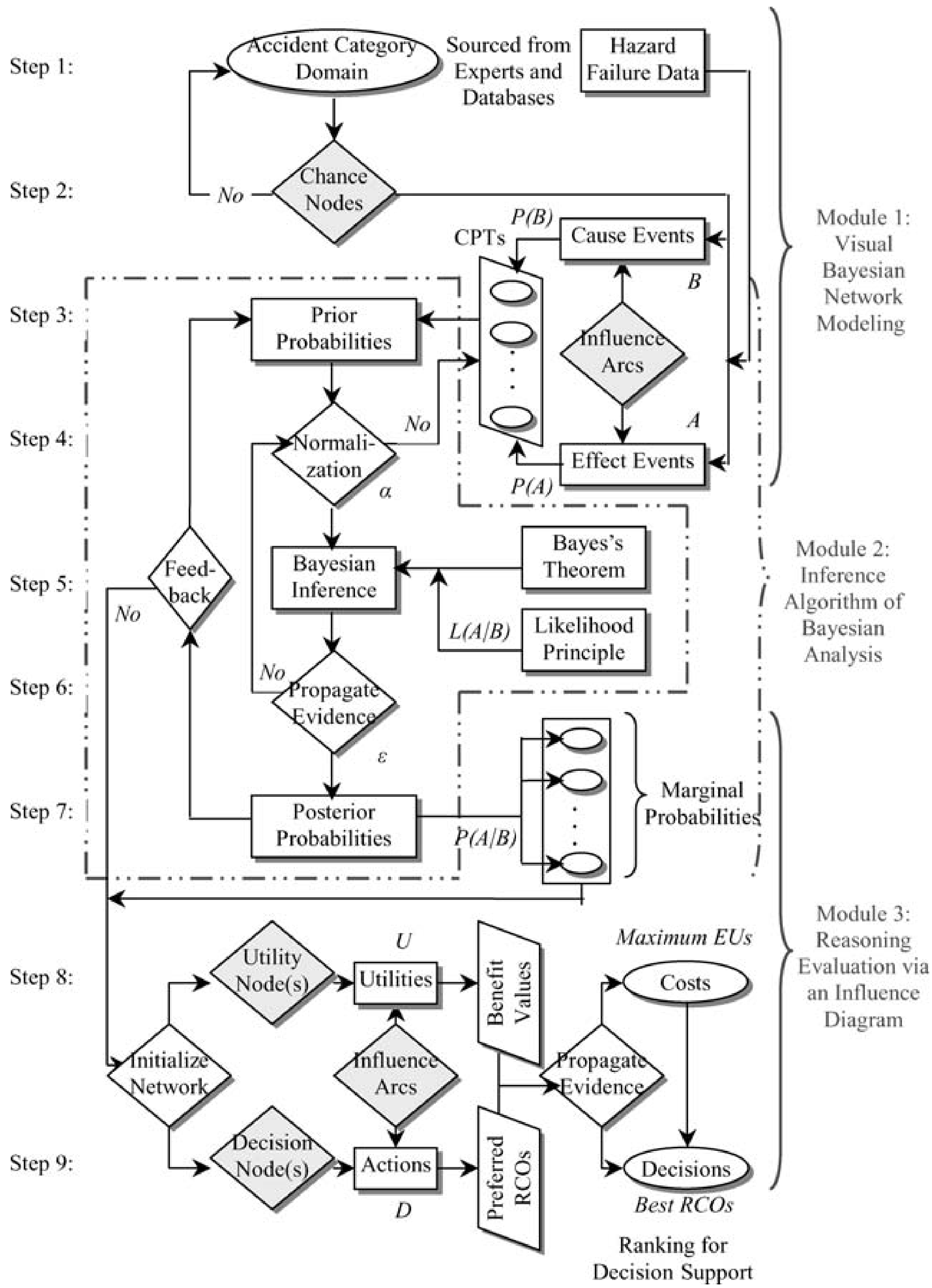


Fig. 2. Flow chart of a proposed BN reasoning framework.

gives the prior probabilities, such as $P(A)$ and $P(B)$, for the nodes. However, nodes without any parents give probabilities that are marginal instead of the conditional ones.

Step 4—Normalization of Probability Values in the CPT: The probability of the marginal and conditional terms being true is nonzero, and becomes 1 after normalization (i.e., the belief values are normalized on a scale from 0 to 1). Thus, the process in this step is to normalize the probability values in every column of CPTs. This normalizing (with an encoded inverse value that gives the normalizing constant, α) has to be done independently for each state of each manifestation across the set of effects.

Step 5—Processing of Data via Bayesian Inference Induction: The Bayesian inference is enabled via the formula: $P(A \mid B) = \alpha L(A \mid B) \times P(A)$, which indicates that the likelihood function, $L(A \mid B)$, is the instrument to pass from prior PD, $P(A)$, to posterior PD, $P(A \mid B)$, via Bayes's theory. $L(A \mid B)$ is induced via LP.

Step 6—Propagation of Evidence: One has to keep in mind that entered evidence propagates in both directions, even though the graph is directed.

Step 7—Generation of Posterior Probabilities: The beliefs computed after evidence is entered to improve the state of knowledge, and thus the prior probability values, are updated by calculating revised probabilities, referred to as the *posterior probabilities*, $P(A \mid B)$. Posterior marginal probabilities, $P(A)$ and $P(B)$, can be obtained via the marginalization process.

If feedback is required due to availability of new data, then the calculated posterior probabilities may become the new prior probabilities for future risk assessment. However, they proceed forward to provide the basis for action.

Step 8—Creation of Decision Node(s) for Preferred RCOs: Initializing the network retracts all findings entered in the risk analysis domain. An ID should be constructed so that one can see exactly which variables (represented by discrete chance nodes) are known at the point of deciding for each decision node. Where the state of a chance node is known at the time of making a decision, one must add a link from the chance node to the decision node. Where the state of a chance node is known before some given decision, and this chance node has impact on another chance node that is also known before the decision, only the last chance node needs to have a link to the decision node. This means that there only needs to be a directed path from a chance node to a decision node if the chance node is known before the decision is made.

Evaluation of the ID is done by setting the value of the decision node to a particular choice of action (i.e., best RCO), and treating the node just as a nature node with a known value that can further influence the values of other nodes.

Step 9—Creation of Utility Node(s) for Values of Achievable Benefits: The action's utility is calculated first by calculating the conditional probabilities for the parents of the utility node using the standard inference algorithm, and then feeding the results to the utility function. The utility figures can be given in terms of property, health, finances, liability, people, environment, public confidence, etc. When propagating, one can follow the EU of choosing each decision in the next decision node in the decision sequence in the node list pane. The best of the RCOs provides the MEU. However, the ranking of the RCOs resulting from the domain case study should aim to be used by decisionmakers at all levels and in a variety of contexts without a requirement of specialist expertise.

8. MARITIME APPLICATION OF REASONING IN BAYESIAN MODELS

To illustrate the universal applicability of BNs and IDs to decision-making problems, it is best to imagine trying to model a situation in which causality plays a role but where an understanding of what is actually going on is incomplete. Thus, things need to be described probabilistically and by inference. Therefore, the demonstration of the modeling and reasoning perspective of this powerful tool is given in the following settings:

- A typical ship evacuation in an accidental risk contribution scenario (a marine case study).
- Authorized vessels to floating, production, storage, and offloading (FPSO) installation collision scenario (an offshore case study).

8.1. Case Study of a Typical Ship Evacuation in an Accidental Risk Contribution Scenario

The safety of people onboard a ship in distress is very much dependent on an effective emergency escape, evacuation, and rescue (EER) operational system (final barrier to avoid fatalities) being in place and being enabled in due time. As the EER system in place

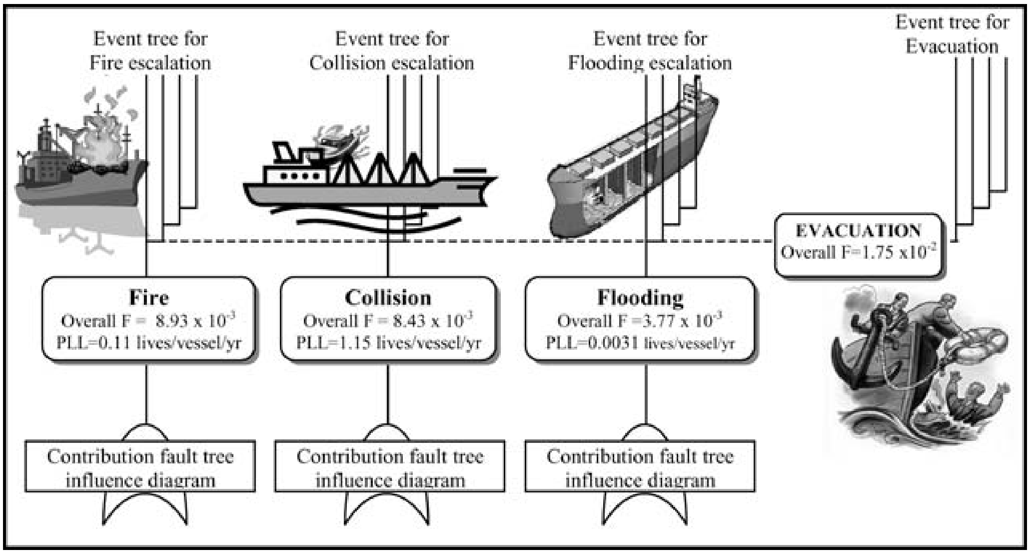


Fig. 3. Risk contribution from major hazards leading to a marine evacuation scenario.

would have to be activated due to the occurrence of some major accident situations, a risk contribution tree (RCT) of the underlying situations may well provide a suitable platform for putting out a BN ship evacuation model. A generic RCT for effecting the evacuation modeling is shown in Fig. 3. It comprises a contribution fault tree and an escalation event tree for the accident categories of fire, collision, and flooding events, together with an evacuation event tree relevant to the accident categories. Each contribution fault tree of the RCT also has the integration of influencing factors (e.g., technical, organizational, and human factors).

The frequency (F) and potential loss of life (PLL) values shown in Fig. 6 are derived from incident databases. Frequency distributions need to be converted into PDs for use in BN, while the PLLs can be applied in cost-effectiveness calculations for use in ID. Since a *failure frequency*, F, in marine assessments is well expressed in terms of per vessel operating year, the overall F values in the RCT can be considered as their *failure rate*, λ, value. If the failure were to follow an exponential distribution, then an equivalent probability value, $P(t)$, for a failure state for the vessel's operational life expectancy, t, is given by:

$$P(t) = 1 - e^{-\lambda t}. \qquad (15)$$

From the case study RCT, this distribution may be used, since it is similar to the discrete Poisson distribution when the occurrence of the event is 0. So, for example, given that a ship has an operational life expectancy of 25 years, evacuation being necessary can be calculated as:

$$P(\text{evacuation necessary}) = 1 - e^{-(1.75\times10^{-2}\times25)} = \underline{\underline{0.355}}.$$

For some typical EER operation (as based on cause-to-effect relationship), a free-fall lifeboat and a rescue boat may be utilized. Thus, a simplified evacuation model to ensure the safety of people onboard a vessel in a distress situation can be represented by the BN model in Fig. 4. Most importantly, the aim of this model and the proceeding analysis is to show how BN can be applied in marine risk assessment while at the same time giving a clearer picture of how a BN model actually works.

To start with, this case study setting has been modeled in a perspective such that "*evacuation being necessary*" does not imply that free-fall lifeboats will not be launched, but that there is a high probability on

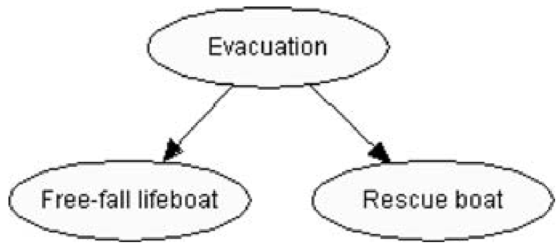


Fig. 4. Simplified BN showing a marine evacuation scenario.

Free-fall lifeboat

Edit Functions View

Evacuation	Unnecessary	Necessary
No launch	0.92	0.04
Launch	0.08	0.96

Fig. 5. CPT for "free-fall lifeboat."

Evacuation

Edit Functions View

Unnecessary	0.645
Necessary	0.355

Fig. 7. CPT for "evacuation."

their launch (or usage). This is modeled in the BN by filling in a CPT for the "*free-fall lifeboat*" node (Fig. 5).

This CPT is actually the *conditional probability* of the variable "*free-fall lifeboat*," given the variable "*evacuation*." The possible values (launch or no launch) for "*free-fall lifeboat*" are shown in the first column. Note that a probability is provided for each combination of events (four in this case). The particular values in this table suggest that the use/launch of free-fall lifeboats is unlikely to increase (8% chance), but once evacuations are necessary, their use is very likely to increase (96% chance). Now, let the use/launch of a rescue boat be considered. To model the uncertainty about whether or not the use of a rescue boat will increase when evacuation is necessary, added to the graph is a new node "*rescue boat*" and an arc from "*evacuation*" to the new node. Although there might not be a great chance that free-fall lifeboats will not be launched, rescue boats may not be quickly responsive in this setting of the evacuation. Therefore, the CPT for "*rescue boat*" (Fig. 6) is different from the one for "*free-fall lifeboat*."

The CPT associated with the node "*evacuation*" is somewhat different in nature. This node has no "parent" node in this example, and consequently, only needs to be assigned a CPT without conditions (Fig. 7).

Determining the probabilities of CPTs is done in several ways. In an instance as this example, it might be a simple case of assigning the probabilities based on the statistical data obtained from a marine incident

Resue boat

Edit Functions View

Evacuation	Unnecessary	Necessary
No launch	0.95	0.35
Launch	0.05	0.65

Fig. 6. CPT for "rescue boat."

database, or from experts with good experience to predict the subjective probabilities.

Having entered the probabilities, the BN can now be used to do various types of analysis. The most important use of BN in this case study is in revising probabilities in the light of actual observations of events (in BN modeling, these are called *evidences* for the maritime BN).

The values of these conditional probabilities can be used to obtain the unconditional probabilities. For example, the unconditional probability that free-fall lifeboats will be launched can be calculated as follows:

$$
\begin{aligned}
&P(\text{free-fall lifeboat launch})\\
&\quad = (P(\text{free-fall lifeboat launch} \mid \text{no-evacuation})\\
&\qquad \times P(\text{no-evacuation})) + (P(\text{free-fall}\\
&\qquad \text{lifeboat launch} \mid \text{evacuation necessary})\\
&\qquad \times P(\text{evacuation necessary}))\\
&\quad = (0.08 \times 0.645) + (0.96 \times 0.355) = \underline{\underline{0.392}}.
\end{aligned}
$$

The rule used here to compute the unconditional probability is called *marginal probability*. Now the unconditional probability that free-fall lifeboats will be launched is known to be 0.392 (i.e., 39.2%).

By running the BN for this evacuation scenario, as can be seen in Fig. 8, *Hugin* gives to the left the *node list pane* and to the right the *modeled network pane*. The *monitor window* placed near the corresponding node in the network pane gives exactly the same as those in the node list pane, and thus they are not always necessary (as they can take up too much space). They are used mainly for nodes that have special interest. As can be seen from the node list pane, as well as that in the monitor window, the unconditional probability that rescue boats will be launched is 26.3%.

Here comes the beauty of BNs. Suppose the launching of free-fall lifeboats is known to increase. In this case, the evidence that "*free-fall lifeboat = launch*" is entered, and then this evidence can be used to determine:

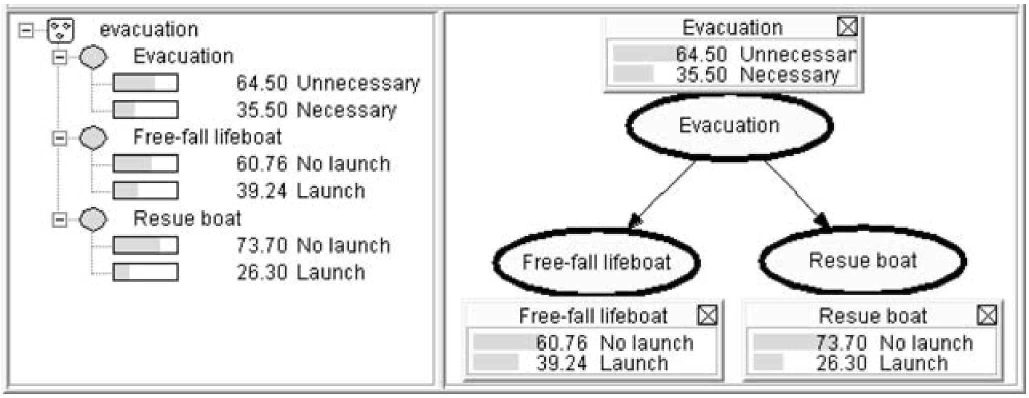


Fig. 8. BN showing results for unconditional probabilities in evacuation scenario.

- The *updated probability* of the necessary evacuation effected.
- The *updated probability* that the use of a rescue boat also increases.

Using Bayes's rule (as presented in Equation (1)), the probability of occurrence for necessary evacuation can be calculated as:

$$P(\text{evacuation necessary} \mid \text{free-fall lifeboat launch})$$
$$= \frac{\left\{ \begin{array}{l} P(\text{free-fall lifeboat launch} \mid \text{evacuation}) \\ \quad \times P(\text{evacuation necessary}) \end{array} \right\}}{P(\text{free-fall lifeboat launch})}$$
$$= 0.96 \times 0.355/0.392 = \underline{\underline{0.869}}.$$

Using *marginal probability*, the probability that there will be a rescue boat launch (see Fig. 9) can be calculated as:

$$\begin{aligned} &P(\text{rescue boat launch}) \\ &\quad = (P(\text{rescue boat launch} \mid \text{no-evacuation}) \\ &\qquad \times P(\text{no-evacuation})) \\ &\qquad + (P(\text{rescue boat launch} \mid \text{evacuation necessary}) \\ &\qquad \times P(\text{evacuation necessary})) \\ &\quad = (0.05 \times 0.131) + (0.65 \times 0.869) = \underline{\underline{0.571}}. \end{aligned}$$

Entering pieces of evidence and using them to update the probabilities in this way is called *propagation*. Fig. 9 shows the results with "*evidence*" node for free-fall lifeboats being launched represented by an *evidence bar* in both the node list pane and in its monitor window in *Hugin*. As would be expected, the probability of occurrence for necessary evacuation increases dramatically to 86.9% when launch of free-fall lifeboats has been observed. This update is due to *diagnosis* (i.e., *bottom-up*) *inference* from the "*free-fall lifeboats*" node to the "*evidence*" node. Furthermore,

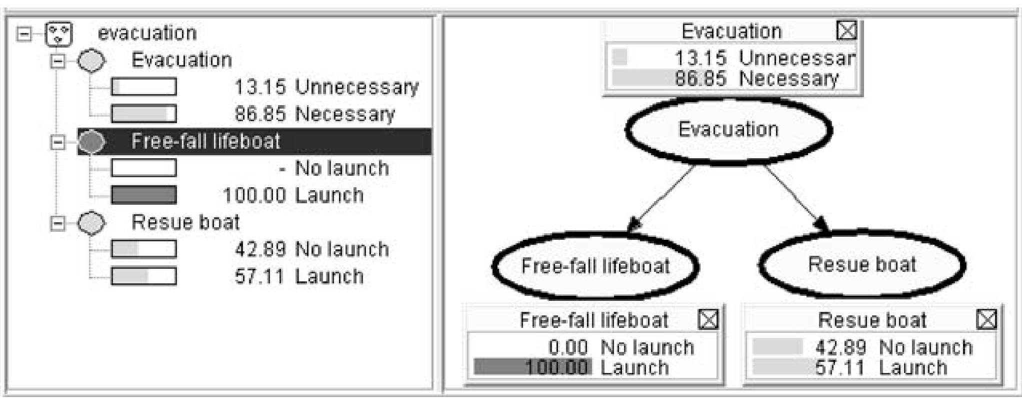


Fig. 9. BN showing propagated results when free-fall lifeboat is launched.

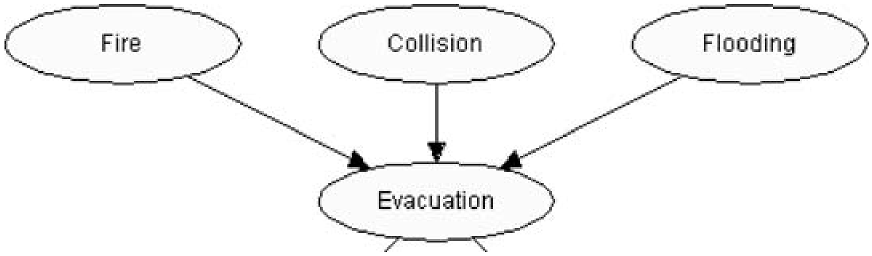


Fig. 10. Fire, collision, and flooding added as parent nodes of evacuation.

Fire

Extinguish...	0.8
Destructive	0.2

Collision

Intact	0.81
Capsize	0.19

Flooding

Discontinuity	0.91
Sinking	0.09

Fig. 11. CPT for each parent node of evacuation.

Evacuation

Flooding	Discontinuity				Sinking			
Collision	Intact		Capsize		Intact		Capsize	
Fire	Extinguish...	Destructive	Extinguish...	Destructive	Extinguish...	Destructive	Extinguish...	Destructive
Unnecess...	0.89	0.35	0.34	0.24	0.16	0.13	0.12	0.01
Necessary	0.11	0.65	0.66	0.76	0.84	0.87	0.88	0.99

Fig. 12. New evacuation CPT reflecting conditional probabilities due to parent nodes.

the updated probability of occurrence for evacuation being necessary results in bringing up the probability for launching of rescue boats to 57.1%, by way of *causal* (i.e., *top-down*) *inference*.

Now, there lies the provision that the major marine accidents of *fire*, *collision,* and *flooding*, which are often variables for external factors, may lead to evacuation. The use of such information has to imply that a *new node* is created and added as *parents* to the evacuation node, for each of these accident categories (Fig. 10).

These new root nodes (i.e., nodes without parents) of evacuation require a CPT without conditions, as they do not have other influences acting on them (Fig. 11).

For the evacuation node, on the other hand, an expanded new CPT is used to reflect the fact that it is now conditional on its three parent nodes (i.e., "*fire*," "*collision*," and "*flooding*"). In other words, the evacuation CPT provides "*P*(*evacuation* | *fire, collision, flooding*)" (see Fig. 12).

Given that in the event of fire or/and flooding *an alarm* will be triggered, a suitable alarm node as *child node* (shown as the highlighted nodes in Fig. 13) can each be linked from the nodes of "*fire*" and "*flooding*," respectively.

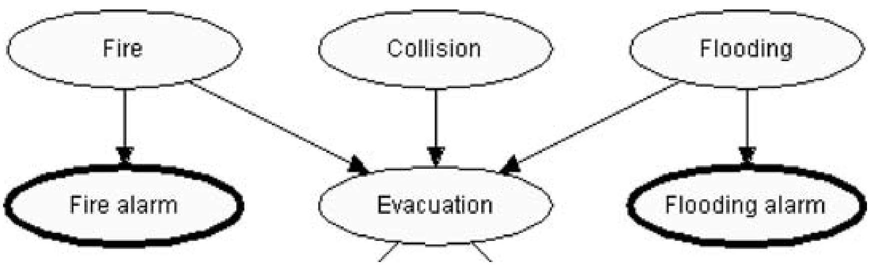


Fig. 13. A suitable alarm added as individual child node to fire and flooding.

Fire alarm

Fire	Extinguishable	Destructive
Not Activated	0.87	0.001
Activated	0.13	0.999

Flooding alarm

Flooding	Discontinuity	Sinking
Not Activated	0.89	0.001
Activated	0.11	0.999

Fig. 14. CPT for individual alarm nodes of fire and flooding.

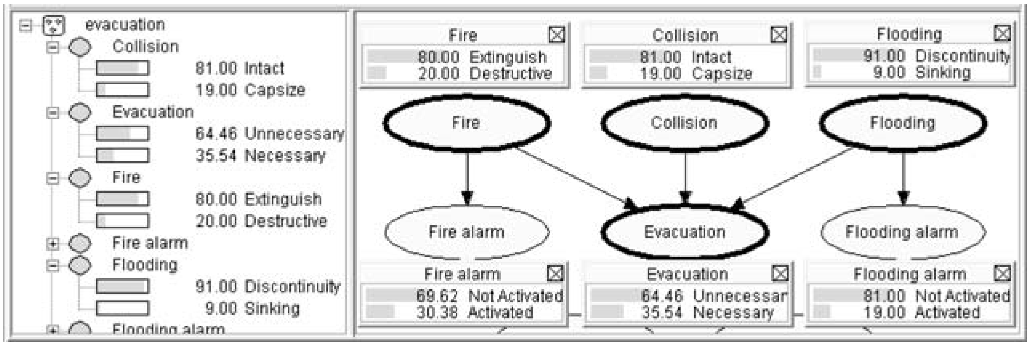


Fig. 15. BN showing marginalized probabilities of evacuation node and its parents.

Since each of the new alarm nodes acts on entirely different accident events, their respective CPTs provide input values of different conditional probabilities (Fig. 14).

Analyzing from the fact that the JPD "*P*(*evacuation, fire, collision, flooding*)" is known, the unconditional probability that evacuation is necessary, "*P*(*evacuation 'necessary'*)" can be given by marginalizing out the "*fire*," "*collision*," and "*flooding*" variables. *Hugin* computes the marginal probability as 35.54% or 0.355 (Fig. 15). Note that *Hugin* also gives the values of 0.304 and 0.19 as the marginal probability of the "*fire alarm*" and "*flooding alarm*," respectively.

In this initialized situation, the root nodes are characterized by their prior probabilities. It is shown in Fig. 15 that the probability of fire being in its destructive state is 0.20, the probability of collision being in its capsize state is 0.19, and the probability of flooding being in its sinking state is 0.09. Suppose it is observed that "*evacuation is necessary*," then this entered evidence increases the belief in all of the possible causes (namely, "*destructive*" for fire, "*capsize*" for collision, and "*sinking*" for flooding) based on *diagnostic inference*. Specifically, applying Bayes's theorem yields a revised probability for fire in destructive state of 0.388 (up from the prior probability of 0.20), a revised probability for collision in capsize state of 0.374 (up from the prior probability of 0.19), and a revised probability for flooding in sinking state of 0.217 (up from the prior probability of 0.09) (Fig. 16). Nonetheless, these revised probabilities are

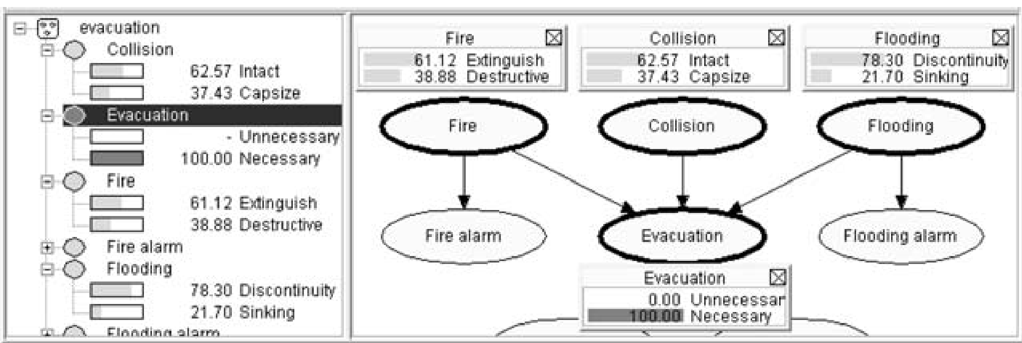


Fig. 16. BN showing propagated results of evacuation evidence to its parent nodes.

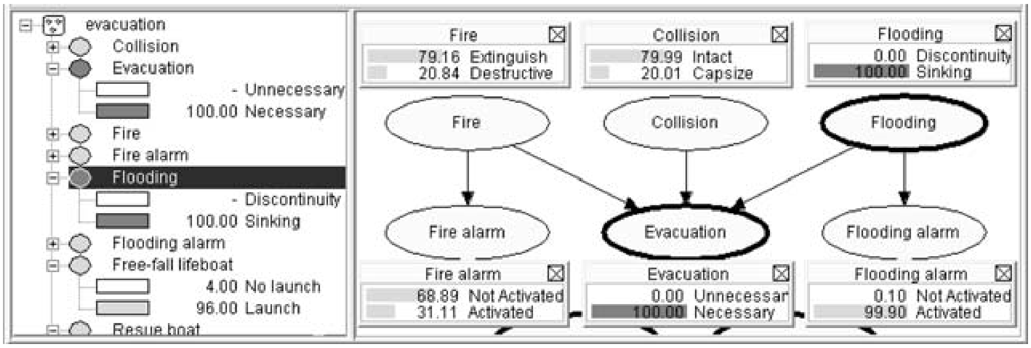


Fig. 17. BN showing propagated results of both evacuation and flooding evidence.

subject to change by the provision of some *additional* observation(s), for example:

- The *additional evidence* firmly on the vessel sinking due to flooding; or
- The *additional evidence* that the fire alarm is activated.

If additional evidence would be firmly on the vessel "*sinking due to flooding*" as the more likely cause, then adding this evidence and applying Bayes's rule would cause the increased probability of "*destruction by fire*" and "*capsize by collision*" to drop to 0.208 and 0.20, respectively (as shown by the monitor windows of Fig. 17), thus "*explaining away*" the "*destruction by fire*" and "*capsize by collision*" as a cause for the "*evacuation being necessary.*" This phenomenon is due to *intercausal* inference.

Conversely, if it is discovered that the fire alarm is activated, then entering this evidence and applying Bayes's rule would yield the revised probabilities of 0.83 for destruction by fire, 0.259 for capsize by collision, and 0.144 for sinking by flooding (as shown by the monitor windows of Fig. 18). Thus, the odds are that the destructive fire, rather than capsize due to collision and sinking due to flooding, has caused the evacuation to be necessary. Once again, it is said that the necessary evacuation has been "*explained away*."

Now, going back to when only evacuation being necessary is observed, the launch of free-fall lifeboats and rescue boats are seen to have a probability of 0.96 and 0.65, respectively (Fig. 19), as induced by *causal inference*. However, when the additional evidence of "*flooding by sinking*" is entered, these respective probabilities remain unchanged (Fig. 20). It is said that the "*evacuation*" node *d-separates* all of its respective parent nodes from each other.

The notion of d-separation (which follows from *human perception*) can also be noticed where only evidence is given for "*flooding by sinking.*" In this case, evacuation being necessary increases from a probability of 0.355 (see Fig. 14) to 0.856 (Fig. 21), but the probability values in the nodes for "*fire*" and "*collision*" stay the same (refer to Fig. 15), as they are not the cause for the increase in probability of the "*evacuation*" being necessary. Thus, the *path* from the "*flooding*" node to these other nodes is *blocked* at the evacuation node. However, the probability values for the launch of free-fall lifeboats being 0.393 and rescue

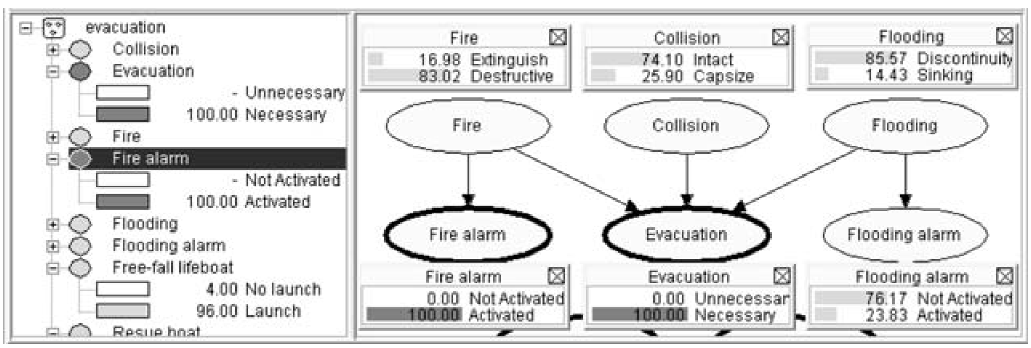


Fig. 18. BN showing propagated results of both evacuation and fire alarm evidence.

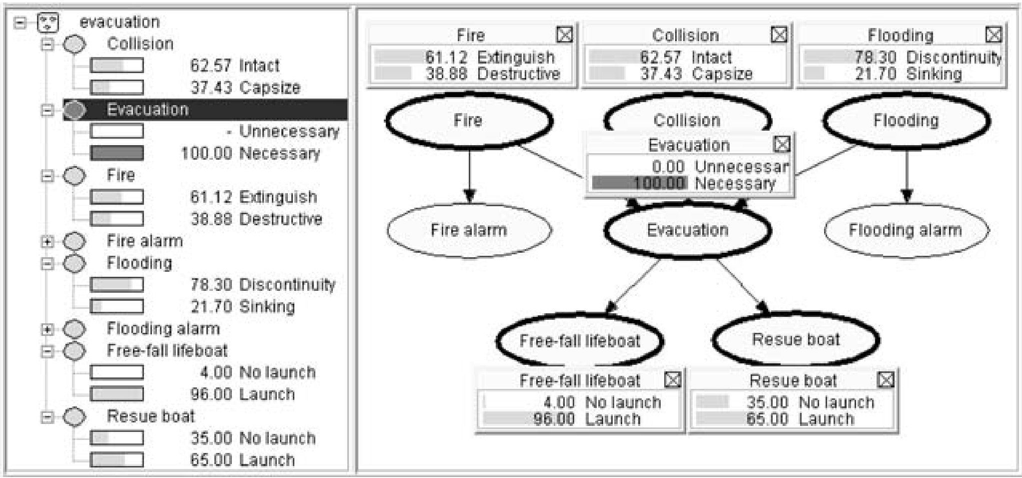


Fig. 19. BN showing evacuation evidence propagation to free-fall lifeboats and rescue boats.

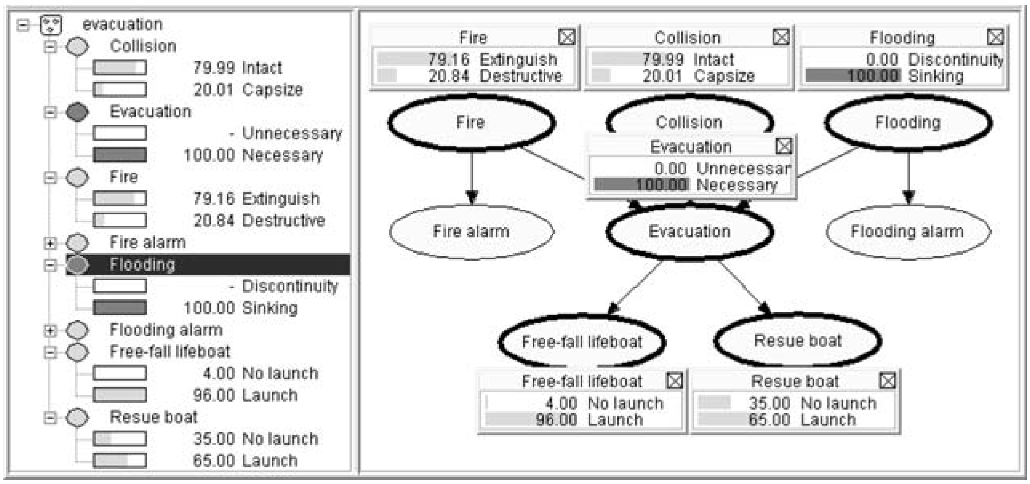


Fig. 20. Flooding and evacuation evidence propagation to lifeboats and rescue boats.

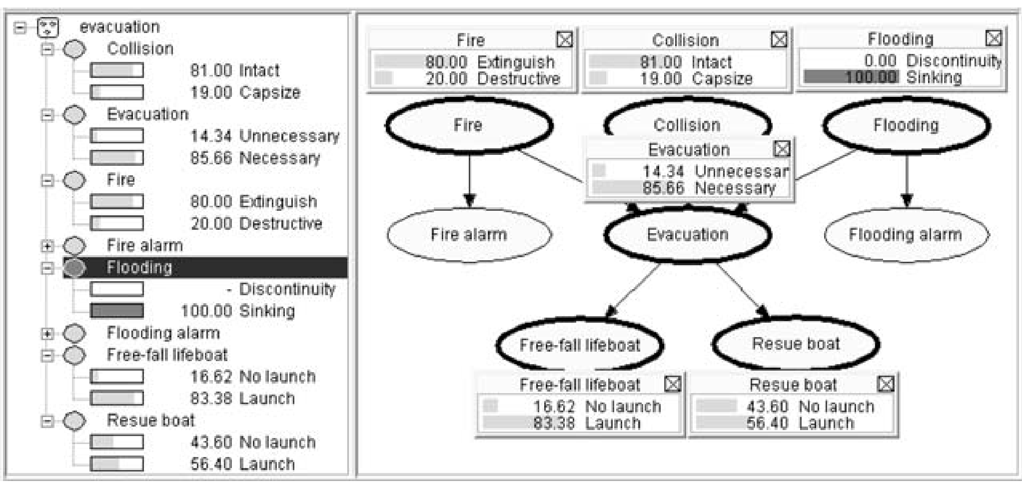


Fig. 21. BN showing evidence of flooding being propagated to evacuation.

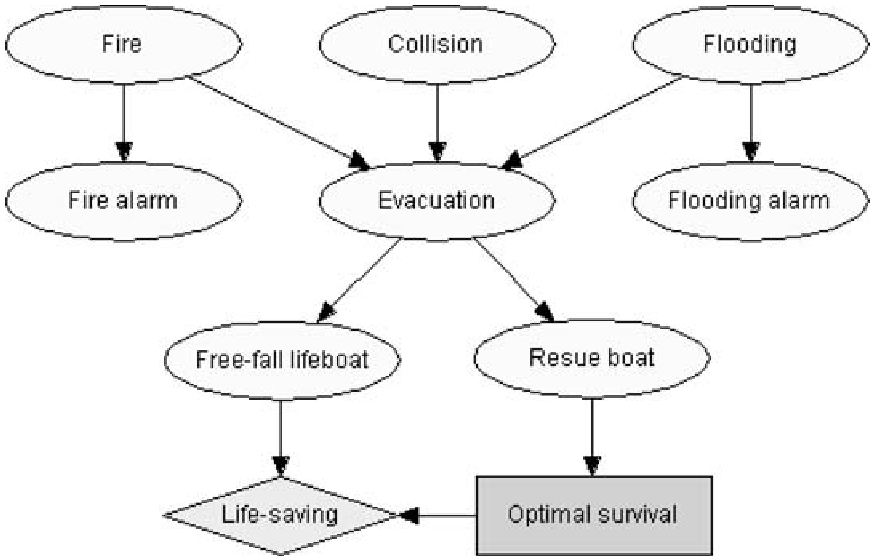
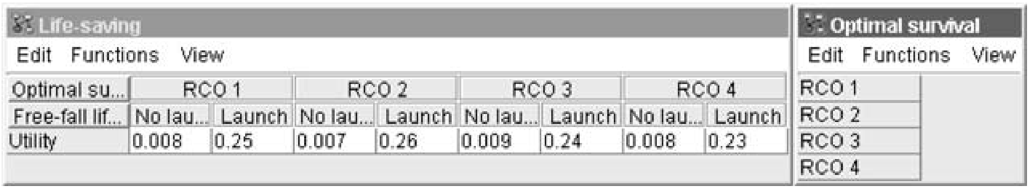


Fig. 22. Simplified ID showing a marine evacuation domain.



Life-saving
Edit Functions View

Optimal su...	RCO 1		RCO 2		RCO 3		RCO 4	
Free-fall lif...	No lau...	Launch	No lau...	Launch	No lau...	Launch	No lau...	Launch
Utility	0.008	0.25	0.007	0.26	0.009	0.24	0.008	0.23

Optimal survival
Edit Functions View

RCO 1
RCO 2
RCO 3
RCO 4

Fig. 23. Encoded inputs in both the nodes of optimal survival and life-saving.

boats being 0.263 (Fig. 8) increase to 0.834 and 0.564, respectively (Fig. 21).

From the analysis so far, although the launch of "*free-fall lifeboat*" and "*rescue boat*" both depend on "*evacuation*" being necessary, "*rescue boat*" launch appears to output a probability value that is less than that of the "*free-fall lifeboat*" launch. The risk analyst has the opportunity to do something about this outcome situation. Thus, a *decision node* that depends upon the rescue boat is added into the model, thereby converting the network into an ID. This new type of node will permit the modeling of an effective decision-support solution that outputs *optimal survival* for those onboard the vessel.

Before the ID is finished, a *utility function*, which gathers information for the potential benefits that come with the different implementation options, and as well, enabling the risk analyst to calculate the EU of the optimal survival, needs to be specified. Given the *outcome state* of "*free-fall lifeboat*," a *value node* of *life-saving*, based on the value of lives saved, is created for specifying these quantitative benefits as a function of the decision. Fig. 22 presents the overall view of this evacuation domain ID.

Where a *formal safety assessment* study has been undertaken for such an evacuation scenario, various RCOs can be identified as decision alternatives based on their cost effectiveness. For the purpose of this case study, the *optimal survival* node has been issued with four hypothetical alternatives, RCO1, RCO2, RCO3, and RCO4, for which the utility value of saving life for the "*launch*" of "*free-fall lifeboat*" is specified as £0.25M, £0.26M, £0.24M, and £0.23M, respectively. The "*no launch*" case, on the other hand, is quantified as £0.008M, £0.007M, £0.009M, and £0.008M, respectively. Fig. 23 shows the quantitative inputs for both the optimal survival decision node and the life-saving utility node (unit is in 10^6 GBP). As seen, the tabular format for the decision node for optimal survival gives just the listing of the entire decision alternatives.

Hugin can then calculate the EU for all of the RCOs as follows:

$$\begin{aligned}\mathrm{EU}(\mathrm{RCO}) = {} & P(\text{no-launch of lifeboat} \mid \mathrm{RCO}) \\ & \times U(\mathrm{RCO}, \text{no-launch of lifeboat}) \\ & + P(\text{launch of lifeboat} \mid \mathrm{RCO}) \\ & \times U(\mathrm{RCO}, \text{launch of lifeboat}).\end{aligned}$$

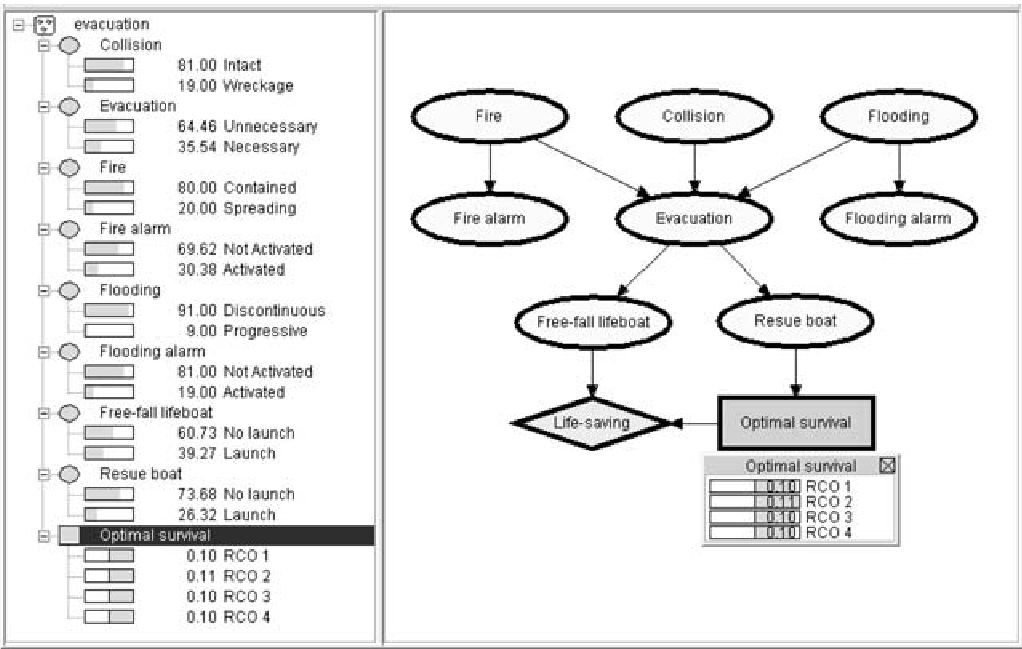


Fig. 24. ID showing initialized values for optimal survival EU.

When no observations are made, the EU values for RCO1, RCO2, RCO3, and RCO4 are assigned with £0.10M, £0.11M, £0.10M, and £0.10M, respectively (Fig. 24). On another note, if an RCO implies large economic benefits with safety implications, it would display a lower net cost of averting a fatality (NetCAF). Thus, *NetCAF* may be used in place of EU to identify which RCOs are justifiable from a commercial or combined commercial and safety point of view.

Once any observation is made, it propagates the evidence by *message passing* and therefore updates the free-fall lifeboat probability. This, in turn, recalculates the EU values for the four decision alternatives. As the best RCOs are those that give the MEUs of optimal survival decision, the RCOs can be ranked accordingly for use in the decision-making process. The MEU is calculated as:

$$\mathrm{MEU(RCO)} = \max_{\mathrm{RCO}} \{\mathrm{EU(RCO1)}, \mathrm{EU(RCO2)}, \mathrm{EU(RCO3)}, \mathrm{EU(RCO4)}\}.$$

In a worst-case scenario, collision might cause damage to the structural integrity of the vessel. As a result, capsize and flooding might result in the sinking of the ship. Since those onboard the vessel need to survive such a disaster, the RCOs for optimal survival are given a *ranking profile* according to their MEU. The MEU order ranking is RCO2 (£0.23M), RCO1 (£0.22M), RCO3 (£0.21M), and RCO4 (£0.20M), as shown in the monitor window in Fig. 25. Thus, the recommendation is for RCO2 and RCO1 to be given top priority with respect to implementation of the optimal survival strategy.

A number of entered evidence circumstances for this model can be investigated. For example, even with the accidental evidence of all root nodes entered, the calculated MEU emerges again with a ranking order of the RCOs as RCO2, RCO1, RCO3, and RCO4, although higher MEU values are reached in this setting (as displayed in the node list pane of Fig. 26).

In the initialized situate, however, it is imperative to determine how "sensitive" the BN ship evacuation model output node results for "*evacuation*," "*free-fall lifeboat*," "*rescue boat*," and "*optimal survival*" are to the input change in variation between the range of lowest and highest possible values that each key event node of "*fire*," "*collision*," or "*flooding*" (as well as any combination of these events) can take. If the model follows the real-world phenomena, then an increase/decrease in the rate or probability at which any of its input event(s) may occur would certainly result in the effect of a relative increase/decrease in the rate or probability of occurrence of its output events.

For example, a *partial sensitivity analysis* for $\pm 20\%$ change to the probability of fire spreading can

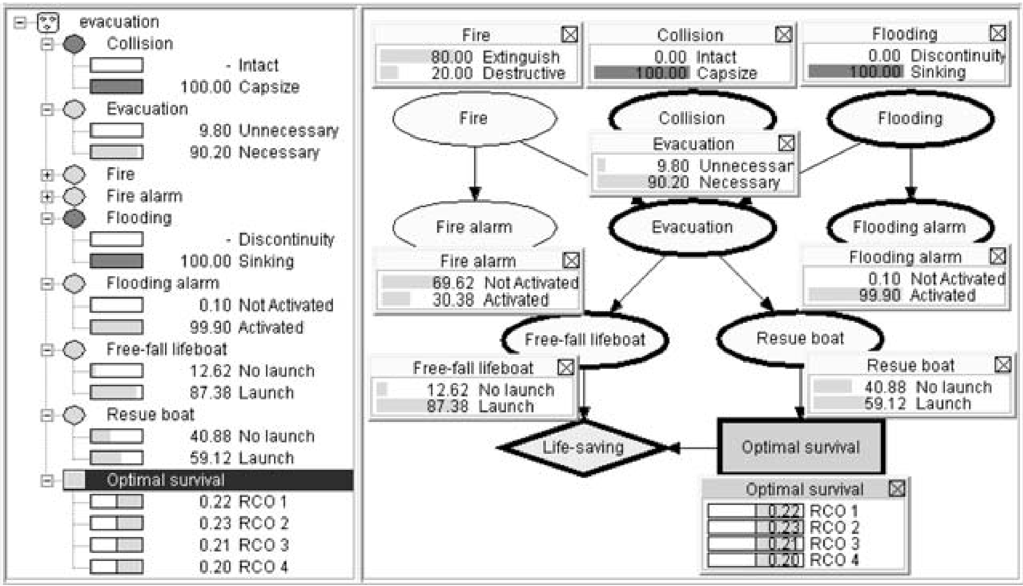


Fig. 25. ID showing propagated results of both collision and flooding.

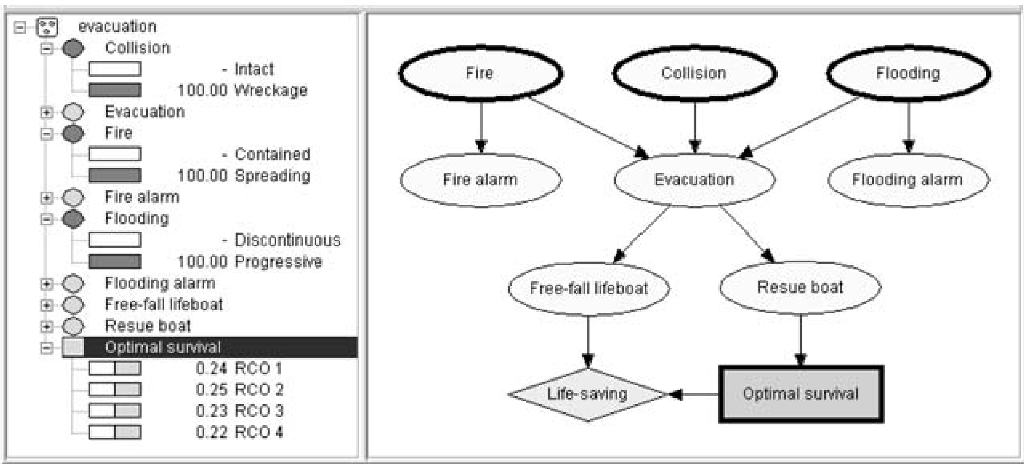


Fig. 26. ID showing optimal survival MEU after entered evidence on all key root nodes.

provide a more realistic setting for which risk analysts and decisionmakers can well determine the response in terms of change in magnitude and direction of the resulting output events. To do this sensitivity analysis, the lowest probability value in the range, which is 0.16 (i.e., −20% of the initial probability of fire spreading value), replaces the initial input value of 0.20; and then using marginal probability, the probabilities of evacuation being necessary, free-fall lifeboat launch, and rescue boat launch are calculated as 0.339 (≈ −4.7% change), 0.378 (≈ −3.7% change), and 0.253 (≈ −3.8% change), respectively (see Fig. 27). Likewise, the MEU for optimal survival becomes £0.10M for RCO1, £0.10M for RCO2, £0.10M for RCO3, and £0.09M for RCO4.

In repeating the sensitivity analysis calculation after substituting the highest probability value in the range, which is 0.24 (i.e., +20% of the initial probability of fire spreading value), the probabilities of evacuation being necessary, free-fall lifeboat launch, and rescue boat launch are calculated as 0.372 (≈ +4.7% change), 0.406 (≈ +3.7% change), and 0.273 (≈ +3.8% change), respectively (see Fig. 28). Similarly, the MEU for optimal survival becomes £0.11M

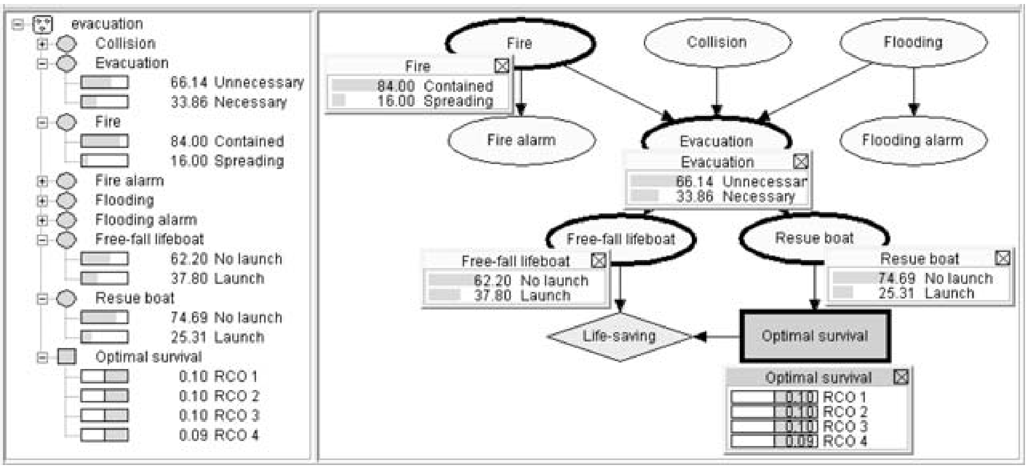


Fig. 27. ID showing model output values for an initialized −20% of *P*(fire spreading).

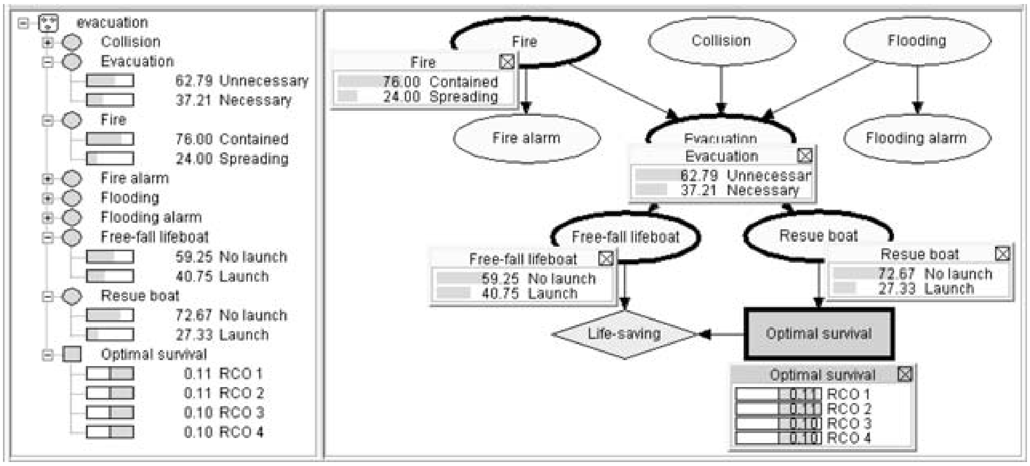


Fig. 28. ID showing model output values for an initialized +20% of *P*(fire spreading).

for RCO1, £0.11M for RCO2, £0.10M for RCO3, and £0.10M for RCO4.

From the sensitivity study, the effects of the ±20% variation in *P*(fire spreading) reveal that this input parameter is a *linear function* with respect to the probability of the evacuation model outputs. Although the decision for optimal survival is sensitive to the state value of *P*(fire spreading), it does not quite reveal the ranking order in the ±20% variation setting.

To establish well the best-ranking order for *P*(fire spreading), a graphical form of the sensitivity analysis may be considered. Based on just varying *P*(fire spreading) through [0, 1], as can be seen in Fig. 29, it is clear that RCO2 gives the best decision alternative while RCO4 gives the worst option to implement. RCO1 appears to overlap with RCO3, but in the region of *P*(fire spreading) equals 0.0 to 0.1 and 0.9 to 1.0, RCO1 can clearly be identified as a definite better option over RCO3. Therefore, the overall decision alternative ranking based on *P*(fire spreading) is given as RCO2, RCO1, RCO3, and RCO4.

8.2. Case Study of Authorized Vessels to FPSO Collision Scenario

To offload oil for shipment to market, a ship-shaped FPSO vessel, being stationed in one location will typically be routinely serviced by supply/standby vessels and shuttle tankers moor at the stern of the FPSO. In a well-known generic scenario, FPSOs can be collided by these ships. They have a risk profile different from fixed platforms and commercial trading

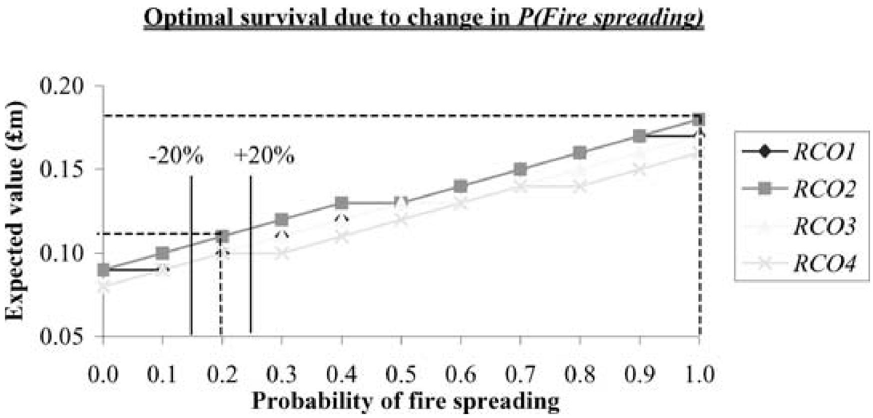


Fig. 29. Effect of varying *P*(fire spreading) for optimal survival ranking.

tankers and, in addition, passing ships also pose a collision risk if an FPSO is close to a sailing route.

The frequency of collision between a shuttle tanker and an installation, or storage unit, is estimated to be 0.0046/year due to failure of the dynamic positioning system. It is assumed that 20% (i.e., 0.0009/year) of shuttle tanker collisions occur after loading operations are complete and the fully loaded vessel is leaving the field. This relatively low percentage is due to the fact that the shuttle tanker is holding and maintaining position, in order to achieve loading, and is aware of the installation's location. In addition, it is usual practice to perform shuttle tanker loading operations at a safe distance from the facility. The remaining 80% (i.e., 0.0037/year) of shuttle tanker collisions are assumed to occur while the tanker is empty and on approach to the facility. Fig. 30 gives the fault tree to estimate frequency of collisions of an FPSO by authorized vessels (Husky Oil, 2000).

The failure of the dynamic positioning system on a maintenance support vessel, causing a collision, is estimated to be 0.0137/year (see Fig. 30).

The evaluation of an FPSO's collision and contact damage risks needs some special technique(s); thus a BN, as shown in Fig. 31, is created in *Hugin* to model this scenario for the FPSO not being able to take measures in avoiding a collision by the authorized vessels maneuvering within close proximity of it. With the ship lifetime and overall production system very conservatively set to 20 years of operation for a lifetime probability in the Bayesian analysis, appropriate probabilities were assigned into the CPT of each node in the model domain. These were based on the failure rates derived from WOAD Statistical

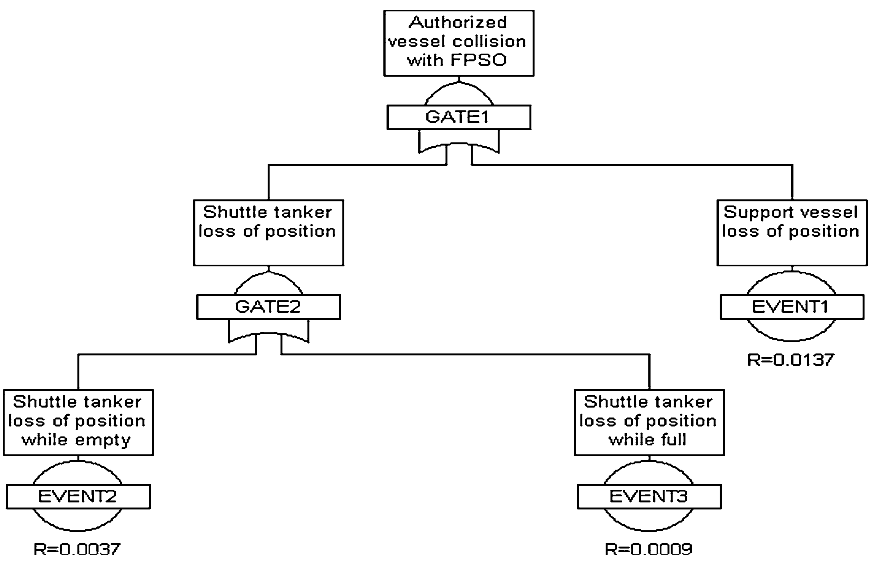


Fig. 30. Fault tree to estimate frequency of collisions of an FPSO by authorized vessels.

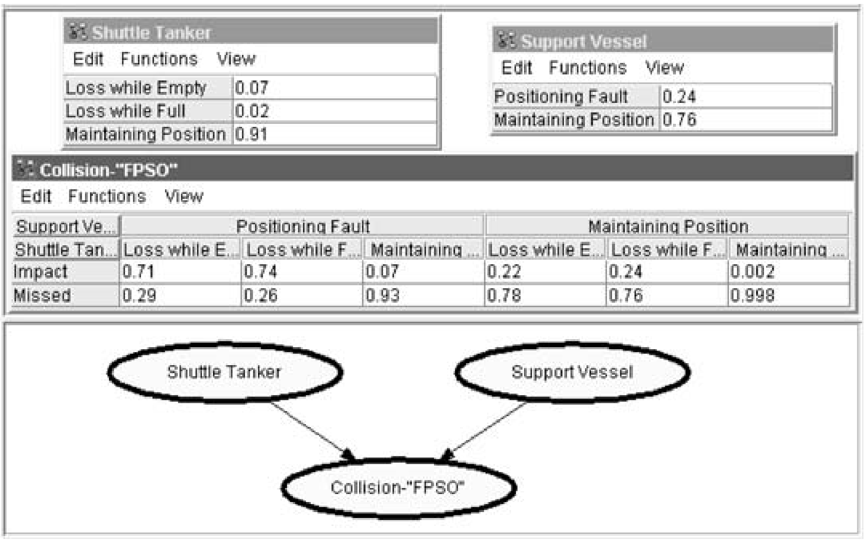


Fig. 31. BN of authorized vessels-FPSO collision scenario with conditional probability tables.

Report (1998) (see Table I) and from the assessments carried out by Husky Oil (2000).

When the net is compiled in "run" mode (Fig. 32), the ship-FPSO collision network window is split into two by a vertical separation, and this gives the initial situation to the left with the node list pane and to the right with the network pane. The probabilities of a node in a certain state are viewed by double clicking it in the node list pane.

Table I. Probability Values from Failure Frequency for Offshore Mobile Units During 1980–1997

Type of Accident	Failure Frequency of Mobile Units (1000 Unit-Years)	Probability (at $t = 20$ Years)
Anchor failure	8.35	0.15
Blowout	10.73	0.19
Capsize	6.56	0.12
Collision	2.78	0.05
Contact	11.53	0.21
Crane accident	4.07	0.08
Explosion	2.78	0.05
Falling load	8.05	0.15
Fire	13.02	0.23
Foundering	5.27	0.10
Grounding	3.18	0.06
Helicopter accident	0.60	0.01
Leakage	3.28	0.06
List	5.86	0.11
Machinery failure	1.39	0.03
Off position	11.53	0.21
Spill/release	9.44	0.17
Structural damage	17.09	0.29
Towing accident	5.86	0.11
Well problem	14.01	0.24
Other	2.48	0.05

To find the probability of the shuttle tanker and the support vessel being in a loss of position failure state, given the information that collision with the FPSO takes place, this fact is entered by double clicking the state "impact" of the Collision-"FPSO" node (Fig. 33). The figure shows the probability of the shuttle tanker being lost while empty to be the most disturbing quantity of the "Shuttle Tanker" node (i.e., 49.75%). Likewise, the "Support Vessel" node now indicates an increase in failure probability to 64.77%.

If it is taken that the shuttle tanker completely (100%) maintains its position, then it can be seen, as in Fig. 34, that the support vessel would have failed drastically in positioning fault (i.e., 91.7%) for there to be a 100% collision impact on the FPSO.

On another note, where collision on the FPSO occurs at either the shuttle tanker being lost while empty (Fig. 35) or while full (Fig. 36), then the Support Vessel node indicates a 50:50 chance of having a positioning fault or maintaining its position.

Evidence identified for nodes being in any state can be added as a node with the links attached from it to these nodes. Some resulting events known to occur due to collision with an FPSO have been identified herein. Some of these, as highlighted in Fig. 37, include spills/release, ignition, explosion, and human injury. Note that the probability values shown in the figure are those for the initial situation in the "run" mode.

When the collision-to-FPSO is set at 100% impact, except for the Ignition node, the failure

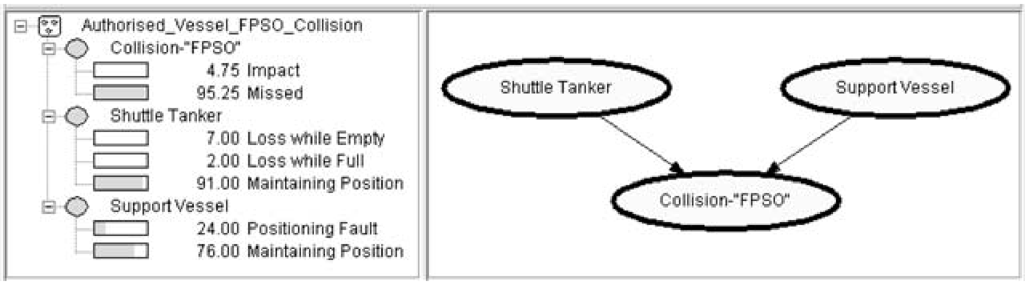


Fig. 32. Initial situation in the BN of authorized vessels-FPSO collision scenario.

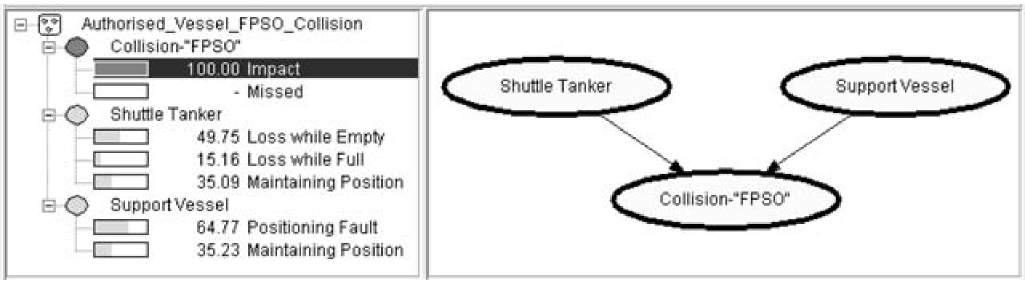


Fig. 33. Probability of impact for Collision-"FPSO" set to 100%.

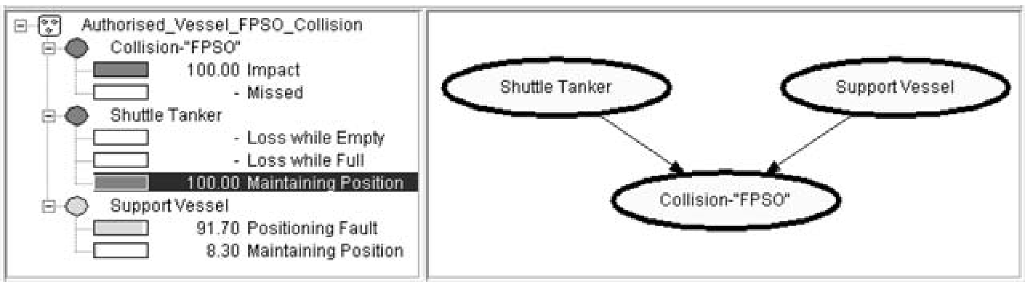


Fig. 34. Collision-"FPSO" impact probability set to 100% in shuttle tanker maintained position.

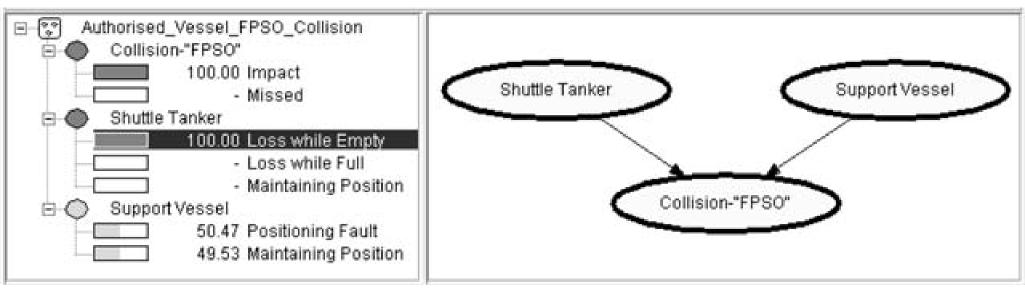


Fig. 35. Collision-"FPSO" impact probability set to 100% in shuttle tanker loss while empty.

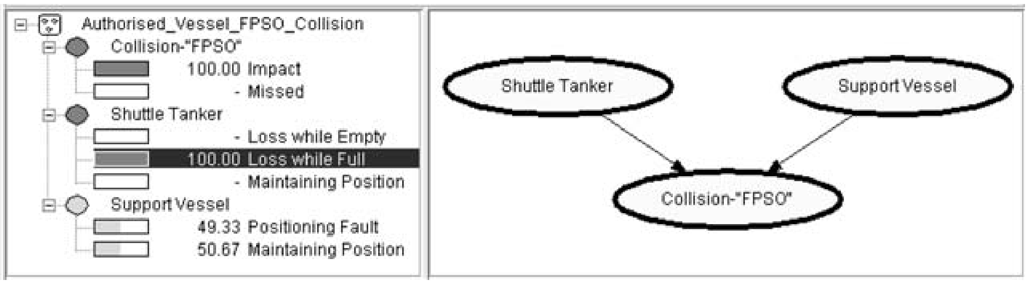


Fig. 36. Collision-"FPSO" impact probability set to 100% in shuttle tanker loss while full.

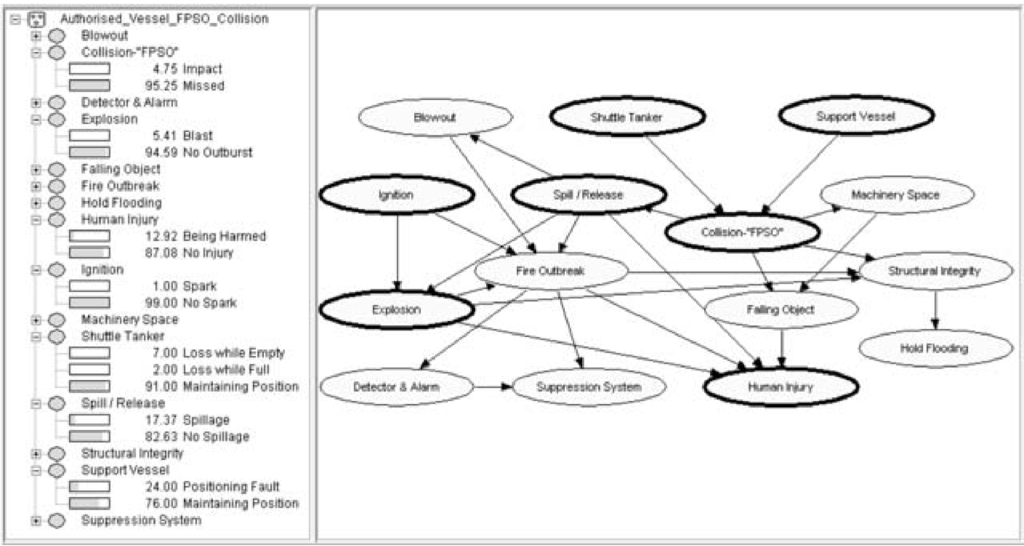


Fig. 37. Addition of evidence and resulting events from the Collision-"FPSO" situation.

probability value of each highlighted node is increased (note from Fig. 38). Those that have significantly increased by a wider margin are especially the Spill/Release node and the Human Injury node. The Ignition node has remained the same in probability value, since it is only a piece of evidence for explosion and fire outbreak, and not a resulting incident of the collision to the FPSO in this scenario.

With the Explosion node set to a failure of 100% blast during a 100% impact on collision with the FPSO, the probability of 96.66% indicates a high amount in certainty for structural damage to happen (Fig. 39). The same can be said for the Human Injury node, which now has a probability value of 84.26%. As such, a great deal of attention will have to be paid to increasing safety for these represented nodes. Thus, the risk analyst and decisionmakers might find it appropriate to consider modeling out an ID for explosion.

Other such pieces of typical evidence as the human element (with states such as "error" and "intervention"), weather condition (with sea states of "calm," "harsh," "adverse," and "severe") (see Fig. 40), electrical/electronic aspects, etc., can be made into new nodes and added to diversify the range of the BN applicability in this scenario.

The scenario settings for this case study can enable a dominant decision in a marine and offshore risk assessment study. Nonetheless, as extensions to the scenario network may lie in the discrepancy of the risk analyst and decisionmakers, the author has chosen to keep the network to an acceptable size. It is best, however, that the risk analyst is aware, in tackling a scenario effectively, of being twisted in the complexities that very large BNs bring.

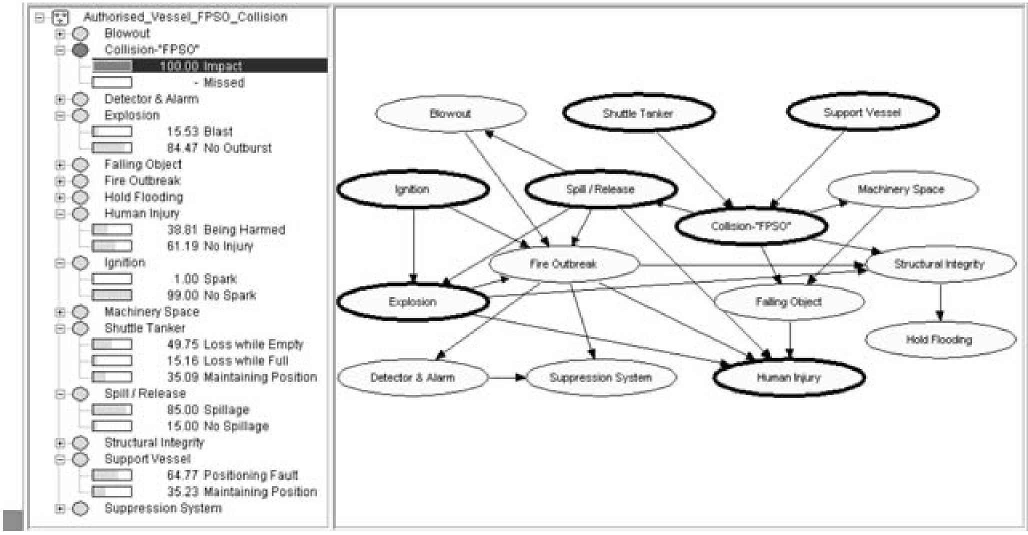


Fig. 38. Situation for resulting events from Collision-“FPSO” impact probability set to 100%.

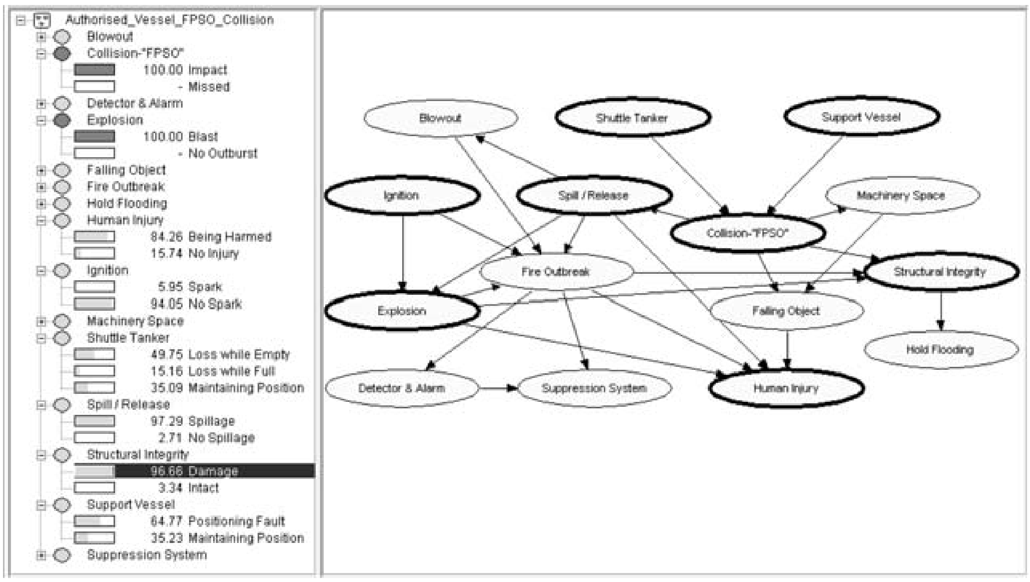


Fig. 39. Situation for resulting events with Collision-“FPSO” and explosion failure set to 100%.

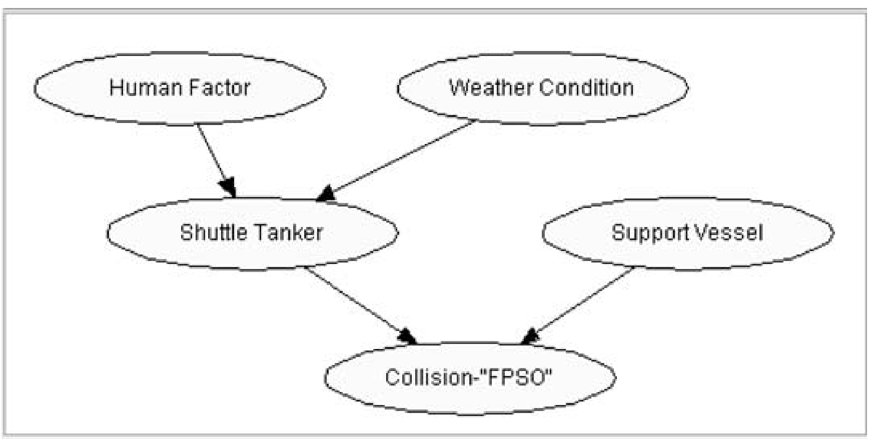


Fig. 40. Some added typical evidence for a shuttle tanker loss of position.

9. BENEFITS AND LIMITATIONS OF BNs

In BNs, each representation possesses particular advantages and disadvantages that make it more, or less, suitable for its intended purpose. These have been recognized and thus outlined in Sections 9.1 and 9.2, respectively.

9.1. Strengths of BNs

The Bayesian framework offers several advantages over alternative modeling approaches. The most important of these advantages are:

- It provides intuitive visual representation with a sound mathematical basis in Bayesian probability that translates into a genuine cause and effect relationship.
- Being probabilistic in its approach, it facilitates a meaningful communication of uncertainty. It is consistent with the risk assessment paradigm, and allows decisions to be made based on expected values.
- It is capable of combining diverse data, expert judgment, and empirical data. By incorporating expert judgment, the method is not paralyzed by a lack of observational data.
- It allows easy updating of prediction and inference in a statistically rigorous manner when observations of model variables are made. Deleting or adding new information does not also require the whole network to be revised.
- The assessment endpoints are chosen so that they are of vital interest to stakeholders and decisionmakers, and can be easily conceived in terms of utility for use in formal decision analysis.

These particular advantages offered by BN make it very useful in situations where uncertainty is unavoidable—Bayesian methods provide a mechanism to model the uncertainty. Thus, such methods can also be used where normal optimization and decision-making techniques are difficult to apply.

9.2. Difficulties of Using BNs

In spite of their remarkable power and potential to address inferential processes, there are some inherent limitations and liabilities to BNs. These drawbacks include the following:

- They cannot easily incorporate unobserved variables, owing to the fact that the size of the internal CPT for a child node can very quickly become quite large.
- There is computational complexity/difficulty (filling in of details of numerical recipe, computer time, convergence monitoring), which is exponential in the number of nodes. These complex models with large numbers of parameters, which are often referred to as nonparametric (NP), become NP-hard in complexity as they approach general multiply-connected networks.
- Likelihood functions are not always solvable analytically (rather, heuristics are extensively used in practice).

The complexity of inference is usually associated with large probabilistic dependencies recorded during inference. However, a large model is preferable to a smaller one only if it provides a sufficiently large improvement of fit to offset the penalty for its additional complexity.

10. CONCLUDING REMARKS

A BN could be used to model the components that affect risk and how they interact. Besides, the graphical nature of a BN makes the model intuitive for users to understand. The process of performing Bayesian updating involves selecting a prior PD, calculating the normalizing constant, formulating the likelihood function, and then calculating the posterior PD. The likelihood function incorporates the objective information, while the prior distribution can include subjective information known about the distributions of the model parameters. Therefore, the posterior distribution incorporates both the *objective* and the *subjective* information into the distributions of the model parameters. Hence, BNs are well suited for modeling maritime safety-critical systems prediction and risk analysis.

The methodology that has been proposed uses BNs to combine evidence from different information sources for a quantitative assessment of a generic scenario. A program tool, such as *Hugin*, can allow the model user to adjust the probabilities of states of nodes based on observed information. The software can also propagate this change through the network, and update the conditional probabilities at each node based on the new information.

As shown in both the ship evacuation and the authorized vessels to FPSO installation collision scenario case studies, by using BNs and a tool such as *Hugin*, it is possible to show all the implications and results of a complex Bayesian argument based on the underlying Bayes's theorem. This theorem is the fundamental principle governing the process of logical Bayesian inference that determines what conclusions can be made with a degree of confidence based on the totality of relevant evidence available. The probabilistic predictions give stakeholders and decisionmakers a realistic appraisal of the chances of achieving desired outcomes. The results from the case studies, as well as other renowned state-of-the-art research work, do indicate that BNs give a sound and transparent approach to modeling marine operational risk. Thus, BN is an integrative model that can be used effectively within the existing decision-making process. BNs can also be expanded to form IDs, which permits rapid development of a practical decision model. The value of IDs as a communication tool has been confirmed. Their use is highly intuitive and they provide a compact alternative to decision trees such that, during review, persons who are not risk analysts are able to interpret the diagrams and propose improvements to the decision model.

ACKNOWLEDGMENTS

This research is partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) through Grant No. GR/S85504 and the Institute of Marine Engineering, Science and Technology (IMarEST) through a Stanley Gray Fellowship.

REFERENCES

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418.

Bolsover, A. J., & Wheeler, M. (1999). Decision-making to treat an explosion hazard. In *Proceedings of the 8th Annual Conference on Safety on Offshore Installations*, *November*. London, UK: ERA Technology.

Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1997). *Expert Systems and Probabilistic Network Models*. New York: Springer-Verlag.

D'Ambrosio, B. (1999). Inference in Bayesian networks. *Artificial Intelligence Magazine*, *20*(2), 21–36.

Edwards, A. W. F. (1992). *Likelihood*. Baltimore: Johns Hopkins University Press.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society, Series A*, *222*, 309–368.

French, S. (1988). *Decision Theory: An Introduction to the Mathematics of Rationality*. New York: John Wiley and Sons.

Gámez, J. A., Serafín, M., & Salmerón, A. (2004). *Advances in Bayesian Networks*. *Series: Studies in Fuzziness and Soft Computing*, *Vol. 146*. London: Springer.

Heckerman, D., Mamdani, A., & Wellman, M. (Eds.). (1995). Real-world applications of Bayesian networks. *Communications of the ACM*, *38*(3), 49–57.

Humphries, P. (1995). *Human Reliability Assessor's Guide*. Human Factors in Reliability Group, Report SRDA-R11. UK: AEA Technology.

Husky Oil. (2000). *White Rose Oilfield Development Application: Assessment of Ship Impact Frequencies*, *Vol. 5*, *Part Two (Concept Safety Assessment), 31 July*. Calgary, Alberta, Canada: Husky Oil Operations Limited.

Jensen, F., Jensen, F. V., & Dittmer, S. L. (1994). From influence diagrams to junction trees. In R. L. Mántaras & D. Poole (Eds.), *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence* (pp. 367–373). San Francisco: Morgan Kaufmann.

Jensen, F. V. (1993). *Introduction to Bayesian Networks: HUGIN*. Aalborg, Denmark: Aalborg University Press.

Jensen, F. V., Lauritzen, S. L., & Olesen, K. G. (1990). Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quarterly*, *4*, 327–352.

Kadie, C. M., Hovel, D., & Horvitz, E. (2001). *MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks*. Microsoft Research Technical Report MSR-TR-2001-67, July. Redmond: Microsoft Corporation.

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society (B)*, *50*, 157–224.

Lindley, D. V. (1970).Bayesian analysis regression problems. In D. L. Meyer & R. O. Collier (Eds.), *Bayesian Statistics* (p. 38). Itasca, IL: F. E. Peacock.

Netica. (2002). *Netica-J Reference Manual—Version 2.21, Java Version of Netica API*. Vancouver, Canada: Norsys Software Corporation.

North, D. W. (1968). A tutorial introduction to decision theory. *IEEE Transaction on Systems Science and Cybernetics*, *4*(3), 105–115.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, *29*(3), 241–288.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems, Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

Russell, S., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*, 2nd ed. NJ: Prentice Hall.

SERENE Consortium. (1999). *SERENE (SafEty and Risk Evaluation using Bayesian Nets): Method Manual*. ESPRIT Project 22187.

Vellido, A., & Lisboa, P. J. G. (2001). An electronic commerce application of the Bayesian framework for MLPs: The effect of marginalization and ARD. *Neural Computing and Applications*, *10*, 3–11.

Von Neumann, J., & Morgenstern, O. (1964). *Theory of Games and Economic Behavior*, 3rd ed. Princeton: Princeton University Press.

Wellman, M. P., & Henrion, M. (1993). Explaining "explaining away". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(3), 287–292.

WOAD. (1998). *Worldwide Offshore Accident Databank (WOAD) Statistical Report*. Høvik, Norway: Veritas Offshore Technology & Services.