A Bayesian Network Approach for Offshore Risk Analysis through Linguistic Variables

Ren J.[1], Wang J.[1*], Jenkinson I.[1], Xu D.L.[2], Yang J.B.[2]
[1]School of Engineering, Liverpool John Moores University
Byrom Street, Liverpool, L3 3AF, UK
[2] Manchester Business School, The University of Manchester
Manchester, M60 1QD, UK

Abstract. This paper presents a new approach for offshore risk analysis that is capable of dealing with linguistic probabilities in Bayesian networks (BNs). In this paper, linguistic probabilities are used to describe occurrence likelihood of hazardous events that may cause possible accidents in offshore operations. In order to use fuzzy information, a f-weighted valuation function is proposed to transform linguistic judgements into crisp probability distributions which can be easily put into a BN to model causal relationships among risk factors. The use of linguistic variables makes it easier for human experts to express their knowledge, and the transformation of linguistic judgements into crisp probabilities can significantly save the cost of computation, modifying and maintaining a BN model. The flexibility of the method allows for multiple forms of information to be used to quantify model relationships, including formally assessed expert opinion when quantitative data are lacking, or when only qualitative or vague statements can be made. The model is a modular representation of uncertain knowledge caused due to randomness, vagueness and ignorance. This makes the risk analysis of offshore engineering systems more functional and easier in many assessment contexts. Specifically, the proposed f-weighted valuation function takes into account not only the dominating values, but also the α-level-values that are ignored by conventional valuation methods. A case study of the collision risk between a Floating Production, Storage and Offloading (FPSO) unit and the authorised vessels due to human elements during operation is used to illustrate the application of the proposed model.

Key Words. Risk analysis, f-weighted valuation function, Bayesian networks, fuzzy number, linguistic probability, offshore engineering systems.

1. Introduction

The key point in offshore risk analysis is how to deal with uncertainty. Uncertainty can be grouped into three categories: randomness, vagueness and ignorance. Randomness is caused due to unpredictable events. It is about the certainty of whether a given element belongs or not to a well-defined set. Classical probability theory is often used to deal with randomness. In offshore risk analysis, some probability theory-based tools such as Fault Tree Analysis (FTA), Decision Table Method (DTM) and Failure Mode & Effects Analysis (FMEA) have been used to solve randomness uncertainties (Wang et al. 1995). Vagueness is mainly caused due to ill-defined concepts in observation or the inaccuracy and poor reliability of instruments used to make observations. Fuzzy set theory can be used to deal with vagueness. Fuzzy reasoning approaches have been developed to deal with problems associated with a high level of vagueness (Ren et al. 2005a; Wang et al. 1995). These include a subjective safety based decision-making method (Wang and Kieran 2000), an evidential reasoning approach, a fuzzy set modelling method (Yang and Xu 2002), and a Dempster-Shafer method for risk modelling and decision-making (Sii et al. 2002). Ignorance is caused due to weak implication, which occurs when an expert is unable to establish a strong correlation between premise and conclusion. In risk analysis,

*: Corresponding author (School of Engineering, Liverpool John Moores University, Liverpool, L3 3AF, UK. Tel: 0151 2312245, Fax: 0151 2312453, Email: j.wang@ljmu.ac.uk)

it is extremely important to find out cause-effect relationships from hazardous factors. For example, a typical question people will ask when investigating an accident is that "what is the cause of the accident?". Bayesian network (BN) is such a powerful cause-effect analysis tool. It is used both for graphically representing the relationships among a set of variables and for dealing with probabilistic variables.

All the three kinds of uncertainties largely exist in maritime risk analysis. In fact, an offshore installation is a complex and expensive engineering structure composed of many systems and is usually unique with its own design/operational characteristics (Wang and Ruxton 1997). Offshore installations need to constantly adopt new approaches, new technologies, new hazardous cargoes, etc., each of which brings a new hazard in one form or another. Therefore, it is essential to reduce the occurrence likelihood of accidents both at the design stage of new facilities and during normal operations, in order to optimise technical and operational solutions. To conduct risk analysis, analysts must deal with fuzzy and incomplete data. In many cases, it may be difficult or even impossible to precisely determine the parameters of a probability distribution for a given event due to lack of evidence or due to the inability of the safety engineer to make firm assessments. One may have to describe occurrence likelihood of events in terms of vague and imprecise descriptors such as "very likely to happen" or "unlikely to happen". These judgements are fuzzy and probabilistic, and hence a novel method is needed to be capable of dealing with such judgements and modelling the safety of offshore engineering systems.

This paper investigates the possibility of using linguistic variables in BNs to provide an alternative means to facilitate offshore risk analysis. The main objective of this research is to propose a novel approach for modelling offshore system safety using BNs. The proposed approach is particularly capable of dealing with linguistic variables. The rest of this paper is organised as follows. Section 2 briefly reviews the major problems and methods in offshore risk analysis. Section 3 gives the background of fuzzy numbers, linguistic variables and BN model. Section 4 discusses f-weighted valuation functions for data transformation. Section 5 proposes an offshore risk analysis flow diagram. Section 6 gives a case study of the collision risk between an FPSO unit and the authorised vessels during operations. Section 7 provides the conclusions of the paper.

2. Major problems and methods in offshore risk analysis

Over the past two decades, a number of serious accidents including the Piper Alpha accident and the Sleipner accident have attracted public concerns to offshore safety and reliability. The studies on how similar accidents may be prevented have been actively carried out at both the national and international levels. Lord Cullen's report on the investigation of the explosion aboard the Piper Alpha platform which claimed the lives of 165 on board was published in November 1990 (UKDE 1990). The report covered a complete range of issues from hardware design and integrity through day-to-day safety management. The inquiry was a stepping-stone to change the safety regime in the offshore industry in the UK. Following Lord Cullen's report, a number of safety regulations have been approved (HSE 1996). Recently, the industrial guidelines on a framework for risk-related decision support have been produced by the UK Offshore

Operators Association (UKOOA) (UKOOA 1999). At the moment, one of the major concerns on the practical application of formal offshore installation safety assessment is associated with the development of integrated and flexible approaches to facilitate its application while human and organizational elements significantly influence the safety of the offshore installation (Wang 2005).

BN is commonly used in establishing the causal relationships among risk elements and estimating the occurrence likelihood of each hazardous event. It is able to replicate the essential features of plausible reasoning in a consistent, efficient and mathematically sound way. Critically it is able to retract belief in a particular case when the basis of that belief is explained away by new evidence (Pearl 1988). BN has been used in many different domains. In recent years, BN has attracted increasing attentions because of the new algorithms developed (Lauritzen and Spiegelhalter 1988; Zhang et al. 2004). BN has the following features:

- It has the ability to incorporate new observations in the network and to predict the influence of possible future observations onto the results obtained (Heckerman and Breese 1996).
- It can not only let users observe the relationships among variables easily, but also give an understandable semantic interpretation to all the parameters in a BN (Myllymaki 2005). This allows users to construct a BN model directly using domain expert knowledge. Furthermore, a BN has both a causal and probabilistic semantics, and thus it provides an ideal representation scheme for combining prior knowledge (which often comes in a causal form) and the historical data.
- It can handle missing and/or incomplete data. This is because the model has the ability to learn the relationships among its nodes and to encode dependencies among all variables (Heckerman 1997).
- It can conduct inference inversely.

Many applications have proven that BN is a powerful technique for reasoning relationships among a number of variables under uncertainty. For example, Hayes (Hayes 1998) applied BN to ecological risk assessment. Kang and Golay (Kang and Golay 1999) applied BN to fault diagnosis in complex nuclear power systems. However, when using BN in offshore risk analysis, there are some difficulties e.g. how to deal with incomplete and vague information that largely exists both at the early system design stage and during normal operations. In the prior research, approximate reasoning approaches have been proposed (Ren et al. 2005a; Sii et al. 2002; Wang et al. 1995) where three fundamental parameters, i.e. *failure rate, consequence severity* and *failure consequence probability,* were used to describe the uncertainties. Due to the high level of uncertainty or the qualitative nature of failure data, risk analysts may often have to use subjective descriptors to describe the above three parameters (Karwowski and Mital 1986; Wang et al. 1995). Furthermore, some quantitative variables are often difficult to be evaluated with precise data. In fuzzy rule-based safety assessment, for instance, the *failure consequence probability* is often directly estimated by expert as a fuzzy number rather than calculated in a more appropriate and rational way e.g. using BN to model and produce the required consequence probabilities. However, conventional BN can only deal with crisp probability and crisp sets. BN is criticised with the utilisation of a probability measure to assess uncertainty. It arguably requires too much precise information in the form of prior and conditional probabilities, and such information is often difficult or impossible to obtain. In particular, in dealing with indirect relationships, even domain experts may find that it is usually difficult to make precise judgments with crisp numbers (i.e. to assign an exact value to the probability that consequences happen given the occurrence of an event). In many circumstances, a verbal expression (e.g. "*very unlikely*") of probabilistic uncertainty may be more appropriate than numerical values. To solve such problems with fuzzy input parameters, it is necessary to develop appropriate novel and flexible risk analysis techniques.

This paper, therefore, proposes a novel and easy-to-use BN method for dealing with uncertainties including randomness, vagueness and ignorance in offshore risk analysis. The idea is to take linguistic variables into account while making good use of conventional BN and its existing algorithms. In doing so, linguistic probabilities are transformed into crisp values by introducing f-weighted valuation function. The transformed crisp values can be put into a conventional BN model supported by mature BN computational techniques. In fact, there are several commercial and research tools designed for BN modelling. Among the most popular of these tools are Hugin (Jenson 1993), Netica (Netica 2002) and *MSBNx* (Kadie et al. 2001). The proposed method provides a potential to use linguistic variables in commercial software tools. The proposed method is also a modular representation of uncertain knowledge and thus can be potentially integrated into the existing offshore safety assessment systems.

3. Fuzzy number, linguistic variables and Bayesian network model

3.1. Fuzzy number

A fuzzy number is a special fuzzy set. It is a way of capturing the impression of a real number by having a fuzzy set. The general definition of a fuzzy number *X* is a fuzzy subset of R (the set of real numbers). If the membership function of X is denoted as μ_X, *X* must meet the following conditions:

(a) The core of *X* is non-empty, i.e. $\exists x \in R$ such that $\mu_X(x) = 1$.

(b) α -cuts of X are all closed, bounded intervals.

(c) It has a *bounded support,* i.e. $\exists N \in R$ such that $\forall x \in R$, if $|x| \geq N$ then $\mu_X(x) = 0$.

Note that an α-cut of a fuzzy number *X* is an interval number X_α that contains all the values of real numbers that have a membership grade in *X* greater than or equal to the specified value of α. This can be written as

$$X_\alpha = [a, b] = \{x \in X | \ \mu_X(x) \geq \alpha\}.$$

3.2. Linguistic variables

Linguistic variables are used to formulate vague descriptions in natural languages in precise mathematical terms. They provide appropriate elements in human knowledge representation. It is important to define the membership function of a linguistic variable to suit different situations. There are different kinds of membership functions. Trapezoid membership function, for example, is one of the most popular. A trapezoidal membership function is normally defined by the parameters *(a, b, c, d)*, where *a* is the membership function's left support with grade equal to *0*, *b* and *c* are the left and right cores with grades equal to *1,* and d is the right support at grade equal to *0*. The function *y = trapezoid (x, (a, b, c, d))* is written to return the membership values corresponding to the defined universe of discourse *x*.

When linguistic variables are used to represent linguistic probabilities in risk analysis of engineering systems, they are expressed by words like *Impossible, Nearly Impossible, Very Unlikely, Unlikely, Even Chance, Likely, Very Likely, Nearly Certain,* and *Certain.* It should be noted that a linguistic term may have different membership functions depending on specific application domains and may vary from case to case and from expert to expert.

3.3. Bayesian network model

A classical BN is a pair N={(V, E), P} where (V, E) represents the nodes and edges of a Directed Acyclic Graph (DAG), respectively, and *P* is a probability distribution over V. Discrete random variables $V = \{X_1; X_2; ...; X_n\}$ are assigned to the nodes, while the edges *E* represent the causal probabilistic relationships among the nodes. Each node in the network is annotated with a Conditional Probability Table (CPT) that represents the conditional probability of the variable given the values of its parents in the graph. The CPT contains, for each possible value of the variables associated to a node, the conditional probabilities with respect to all the combination of values of the variables associated to the parent nodes. For nodes that have no parents, the corresponding table will simply contain the prior probabilities for that variable. The principles behind BN are Bayesian statistics concentrating on how probabilities are affected by both prior and posterior knowledge.

Inference in BN generally targets the calculation of some probabilities of interest. The inference algorithms are based on the following four equations:

Conditional independence

$$P(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} P(X_i \mid Parents(X_i)) \tag{1}$$

Joint probability

$$P(Y = y_j, X = x_i) = P(X = x_i) \cdot P(Y = y_j \mid X = x_i) \tag{2}$$

Marginalization rule

$$P(Y = y_j) = \sum_i P(X = x_i) \cdot P(Y = y_j \mid X = x_i) \tag{3}$$

Bayesian rule

$$P(X = x_i \mid Y = y_j) = \frac{P(X = x_i) \cdot P(Y = y_j \mid X = x_i)}{P(Y = y_j)} \tag{4}$$

In order to use linguistic terms and fuzzy numbers in the above Bayesian rules, a fuzzy-to-crisp value transformation method is thus needed.

4. F-weighted valuation method to transform linguistic/fuzzy probabilities into crisp values

BN modelling and propagation algorithms are based on crisp prior/posterior probabilities. In practice, however, precise probabilities are difficult to be estimated and much data collected from domain experts is often in the form of linguistic terms. Usually a linguistic term can be represented by a fuzzy membership function. In such cases, before the BN modelling and associated algorithms are applied, it is necessary to transform the linguistic probability in the form of linguistic terms into a crisp value.

Several tools are available for fuzzy-to-crisp transformation such as Maximum Transformation Technique (MTT), Centroid Defuzzification Technique (CDT) and Weighted Average Technique (WAT) (Ross 1995). Many existing techniques are suffering from information loss and being sensitive to single information that dominates the fuzzy set during the process that transforms a fuzzy number into a crisp value. This paper adopts f-weighted valuation function (Detynieckim and Yager 2000; Yager 1981) to decrease the degree of losing information and make the analysis results more reasonable and reliable.

A generalized formulation for a class of valuations is as follows:

$$Val(F) = \frac{\int_0^1 Average(F_\alpha) \times f(\alpha) d\alpha}{\int_0^1 f(\alpha) d\alpha} \tag{5}$$

where $Val(F)$ is a crisp value transformed from fuzzy membership function *F*; $F_\alpha = \{x \mid F(x) \geq \alpha\}$ is an α-level set of *F*; $Average(F_\alpha)$ is the average of the elements in the α-level set. f is defined as f-weighted valuation function.

In Equation (5), the choice of the actual function $Val(F)$ is subjective, it depends upon how analysts put attention to the fuzzy membership function *F,* that is, attention should be put more in the core area or in the support area. This provides flexibilities to analysts when they are dealing with real world problems.

When *F* takes the form of trapezoidal fuzzy set, for instance, it will have the membership function:

$$F(x,(a,b,c,d)) = Trapezoid(x,(a,b,c,d)) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ 1 & b < x \leq c \\ \frac{d-x}{d-c} & c < x \leq d \\ 0 & x > d \end{cases}$$

For the case of trapezoidal fuzzy set, the average of the elements in the α-level set can be computed as follows (Yager 1999):

$$Average(F_\alpha) = \frac{u_\alpha + v_\alpha}{2}$$

where u_α and v_α are the horizontal axis values of intersection points between α-cut line and the left-hand side and right-hand side of the trapezoidal fuzzy set, respectively. They are calculated as follows:

$u_\alpha = (b-a)\times\alpha + a$ and $v_\alpha = d-(d-c)\times\alpha$

Then Equation (5) becomes:

$$Val(F(x,(a,b,c,d))) = \frac{\int_0^1 \frac{u_\alpha + v_\alpha}{2}\times f(\alpha)d\alpha}{\int_0^1 f(\alpha)d\alpha} = \frac{\frac{1}{2}\int_0^1 [(b+c)\times\alpha + (1-\alpha)\times(a+d)]\times f(\alpha)d\alpha}{\int_0^1 f(\alpha)d\alpha} \qquad (6)$$

It has been proved that the crisp value given by Equation (6) is a weighted-mean of the average of the core (left and right core) and the average of the support (left and right support). In fact, higher α-level sets are closer to the core of trapezoidal fuzzy set e.g. the α-level set equals the core when α =1. The core of the trapezoidal membership function means the most likely value of the function.

The selection of f-weighted valuation function is dependent on the type of fuzzy membership function F takes. Without loss of generality, let f-weighted valuation function $f(\alpha) = \alpha^q$ where q acts as a control factor in order to smoothen the transformation output. Equation (6) becomes:

$$Val(F(x,(a,b,c,d))) = \frac{\frac{1}{2}\int_0^1 [(b+c)\times\alpha + (1-\alpha)\times(a+d)]\times\alpha^q d\alpha}{\int_0^1 \alpha^q d\alpha} = \left(\frac{b+c}{2}\times\frac{q+1}{q+2}\right) + \left(\frac{a+d}{2}\times\frac{1}{q+2}\right) \qquad (7)$$

Equation (7) is further explained as follows:

- If q = 0, Equation (7) becomes:

$$Val(F(x,(a,b,c,d))) = \frac{\left(\frac{b+c}{2}\right) + \left(\frac{a+d}{2}\right)}{2} \qquad (8)$$

  This shows that the transformed value is between the middle point of the core and the middle point of the support.

- If q increases from 0 to $+\infty$, Equation (7) becomes:

$$Val(F(x,(a,b,c,d))) = \frac{b+c}{2} \quad (9)$$

This shows that the transformed value is exactly at the middle point of the core.

Using different q values in f-weighted valuation function, experts are able to adjust subjective parameter values to make their judgements more rational. Without loss of generality, this paper uses two q values (q=0 and q=$+\infty$), i.e. Equations (8) and (9) will be used in a case study in Section 6.

5. BN-based risk analysis flow diagram

To conduct offshore risk analysis, a uniformed flow diagram is essential. In such a flow diagram, offshore system hazards identification and BN inference process can be dealt with in an integrated manner. The flow diagram is developed to include all significant processes in BN inference and thus provide a basis for further development of a generic offshore risk analysis framework. Based on the literature review, the flow diagram is proposed and depicted in Figure 1. It outlines the necessary steps required for risk analysis in a holistic way supported by BN. It acts as a guideline in carrying out cause-and-effect inference based on subjective judgements. The proposed flow diagram consists of two major components. Particularly, in the BN module the integration of the fuzzy-to-crisp data transformation function and the Bayesian inference function makes the module capable of dealing with uncertainty that is mainly caused by randomness and vagueness. The major steps of the components used in the flow diagram are outlined as follows:

Component 1: Problem definition & potential failure identification

- Identify all anticipated causes/factors to potential failures of an offshore engineering system. Particular attention must be paid to causal relationships among those anticipated causes/factors.
- Define variables (nodes) to represent the identified potential failures.

Component 2: BN module

- Construct a BN model from a generated list of the identified hazards specific to the problem under investigation.
- Select the types of fuzzy membership function used to delineate variables (nodes), and provide interpretation for each fuzzy set of each variable (node).
- Integrate domain specific knowledge obtained from available data included in regulatory rules and standards, databases and data networks, and, if necessary, also from simulations and experiments into the network structure.
- Specify states and assign input values for the fuzzy conditional probability table (FCPT) of each variable (node).
- Transform linguistic/fuzzy data into crisp value.
- Update the values of all nodes by calculating posterior probabilities via the four Bayesian equations (i.e. Equations (1), (2), (3) and (4)) when new information is available; conduct sensitivity analysis.
- Interpret the causal relationships and provide the generated results for approximate reasoning based safety synthesis.

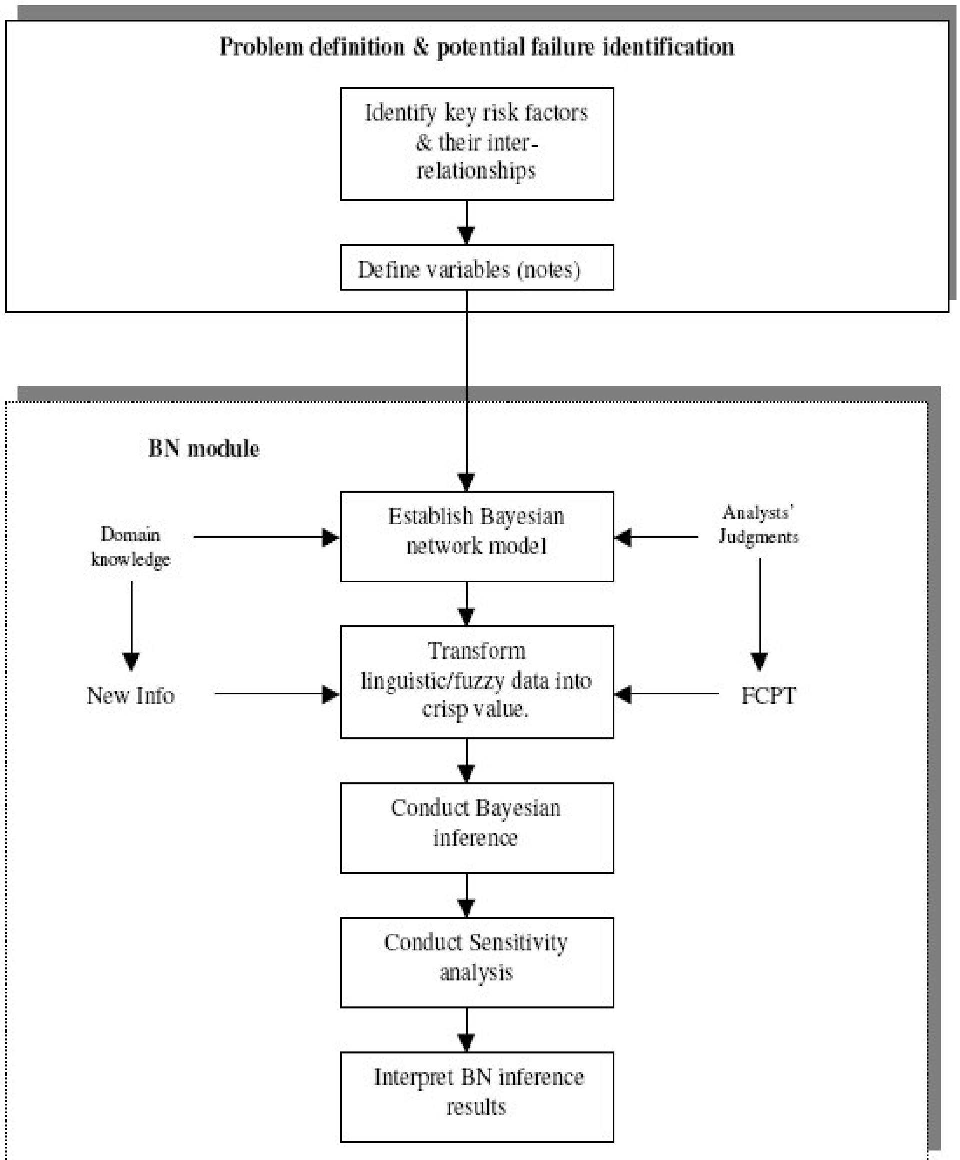

Figure1: Flow diagram of the proposed offshore risk analysis

6. Case study: a collision risk analysis between an FPSO unit and the authorised vessels

Human and organisational errors are the main causes of severe incidents such as the Piper Alpha tragedy. According to the P&I Club studies into accidents, approximately 80% of maritime accidents are attributable to human errors (Mitchell and Bright 1995). Understanding human and organisational errors and system failures is particularly important with respect to offshore installations. In this section, a case study of the collision risk analysis between an FPSO unit and the authorised vessels due to human errors during a tandem offloading operation is carried out to illustrate the application of the proposed methodology. The analysis process follows the diagram proposed in Section 5 (see Figure 1).

6.1. General description and BN model establishment

An FPSO unit is one of the most popular floating systems used by the offshore oil and gas industry. In the UK, crude oil from an FPSO is normally transported to shore using shuttle tankers specially designed for dealing with the harsh weather conditions. Shuttle tankers equipped with a bow-loading system are connected to an FPSO or storage facilities by mooring hawser and loading hose through which cargo is offloaded. Tandem loading/offloading is a complex marine operation. It is with high risk due to the close proximity required between the two large vessels. In addition, the FPSO unit is also routinely serviced by support vessels. During the operation of service, the support vessels could collide with the FPSO unit due to faulty positioning. The consequence of the collision varies from minor contact to severe incidents that may cause personnel injury/loss, environment pollution and/or damage to the vessel. For demonstration purposes, this case study considers six factors: Human Errors (*X*), Adverse Weather (*Y*), Shuttle Tanker Position (*Z*), Malfunction of Support Vessel (*Y*), Collision with an FPSO unit (*U*) and Personnel Injury/Loss (*W*). The causal relationships among those six factors are addressed in a way that human elements or weather conditions may cause the shuttle tanker or the support vessel to be in a faulty position. Such a faulty position of the shuttle tanker or the support vessel may cause collision with the FPSO, and thus may cause personnel injury/loss. The causal relationships are demonstrated in Figure 2 (Ren et al. 2005b). As can be seen in Figure 2, the six nodes are organised by the acyclic arrows that represent the causal relationships among them. One of the most interesting questions is to find out that if there is a personnel injury/loss observed, then in what possibility it is caused by human errors.

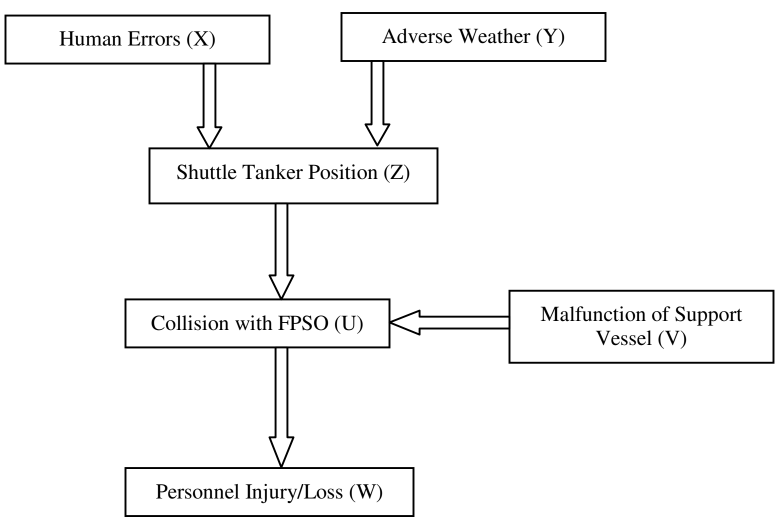

Figure 2: The Bayesian network model of the collision risk analysis between an FPSO unit and the authorised vessels

6.2. Linguistic term definitions

Domain experts and analysts can be invited to define probability likelihood using linguistic terms. Table 1 describes the linguistic labels and their associated meanings of the linguistic variables. As can be seen in the table, for each linguistic term there is a range of probability values to describe it. The meaning of each linguistic variable is explained in the second column of Table 1. It should be noted that a linguistic term may have more than one membership functions e.g. "*Very Unlikely*" may have membership functions (0.19, 0.20, 0.23), (0.22, 0.25, 0.26) or (0.24, 0.25, 0.28). This is because when domain experts assess the occurrence likelihood of three separate possible events, they may use one term to describe them e.g. "*Very Unlikely*". However, if they can tell the difference between the occurrence likelihood of the three possible events, then appropriate membership functions may be assigned to term "*Very Unlikely*". This provides accuracy to linguistic variables as well as keeps the linguistic meanings to fuzzy membership functions.

It is worth noting that the definitions of linguistic probability sets are made to demonstrate the implementation process of the risk analysis approach, though other types of the linguistic labels and their associated meanings of the linguistic variables can also be used, depending on specific circumstances. In solving real world problems, risk analysts and decision makers must be consulted when there is a need to redefine and estimate the linguistic probability sets. Table 1 is just an example used in the demonstration case study of this paper.

6.3. Linguistic/fuzzy prior and conditional probabilities

Domain experts were asked to give judgments about the linguistic/fuzzy probabilities regarding all the nodes in Figure 2. Tables 2, 3 and 5 give fuzzy prior probabilities of nodes *X, Y* and *V*, respectively. Tables 4, 6, and 7 are the fuzzy conditional probabilities of nodes *Z, U* and *W*, respectively. As shown in Table 2, there are two possible values (x_1 or x_2) for "Human Errors (*X*)". If *X* is true (x_1), it means that the errors caused by human elements take place. Domain experts used linguistic description to address the inherent uncertainty associated with the data and information available while carrying out the assessment on the "Human Errors (*X*)" node. In the case study, for example, domain experts assessed the occurrence likelihood of "Human Errors (*X*)" as "*Unlikely*". Triangular fuzzy number (0.39, 0.40, 0.43) is used to describe the linguistic term "*Unlikely*" as shown in Table 2. A triangular fuzzy set is a special trapezoidal fuzzy set when the core set of the trapezoidal fuzzy set takes the form of a single point. The most likely value is 0.4 while 0.39 and 0.43 are the lower and upper least likely values respectively.

Similarly in Table 3, there are two possible values (y_1 or y_2) for node *Y* ("Adverse Weather"). $Y=y_1$ means that "bad weather condition" is observed. As can be seen in Table 3, a triangular fuzzy number (0.49, 0.50, 0.51) is used to describe the linguistic term "*Even Chance*" which gives the occurrence likelihood of "Adverse Weather".

Table 4 gives the conditional fuzzy probabilities of variable *Z* ("Shutter Tanker") given the states of nodes "Adverse Weather (*Y*)" and "Human Errors (*X*)". In Table 4, a fuzzy probability is provided for each possible combination of states of nodes *X* and *Y* (eight in this case). The fuzzy probability value ($Z=z_1$) under condition of x_2 and y_2, for example, is shown in the fifth row and third column. The particular value suggests that the faulty position of the shuttle tanker is "*Nearly Impossible*" to happen with fuzzy probability (0.04, 0.05, 0.08) given $X= x_2$ and $Y= y_2$. However if human operational error (x_1) has happened under bad weather condition (y_1), the occurrence likelihood of faulty position of the shuttle tanker is increased to "*Quite Likely*" with fuzzy probability (0.64, 0.65, 0.68) (shown in the second row and third column of Table 4).

Table 1 Linguistic labels and corresponding meanings with examples of possible fuzzy membership functions

Linguistic label	Meaning	Examples of possible fuzzy membership functions
Impossible	Events never happen.	0.0
Nearly Impossible	The occurrence likelihood of possible events is *nearly impossible* (extremely unlikely to exist during operations).	(0.001,0.002,0.005) (0.04,0.05,0.08)
Very Unlikely	The occurrence likelihood of possible *events* is highly unlikely (highly unlikely to exist during operations).	(0.19,0.20,0.23) (0.22,0.25, 0.26) (0.24,0.25, 0.28)
Unlikely	The occurrence likelihood of possible *event*s is unlikely but possible (improbable to exist even on rare occasions during operations).	(0.34,0.35,0.38) (0.39,0.40,0.43)
Even Chance	The occurrence likelihood of possible *events* is *even chance* (likely to exist on rare occasions during operations).	(0.49,0.50,0.51)
Likely	It is likely that *events* occur (i.e. exist from time to time during operations, possibly caused by a potential procedural weakness in design or operation).	(0.57,0.60,0.61) (0.62,0.65,0.66)
Very Likely	It is highly likely that *events* occur (i.e. often exist somewhere during operations due to a highly likely potential hazardous situation or procedural weakness in design and/or operation).	(0.72,0.75,0.76) (0.74,0.75,0.78) (0.77,0.80,0.81)
Nearly Certain	*Events always* happen (i.e. likely to exist repeatedly during operations due to an anticipated potential procedural weakness in design and operation).	(0.92,0.95,0.96) (0.995,0.998,0.999)
Certain	*Events definitely* happen.	1.0

Table 2 Fuzzy prior probability $P_f(X)$

X	$P_f(X)$
x_1	(0.39,**0.4**,0.43), <u>0.405</u>
x_2	(0.57,**0.6**,0.61), <u>0.595</u>

Table 3 Fuzzy prior probability $P_f(Y)$

Y	$P_f(Y)$
y_1	(0.49,**0.5**,0.51), <u>0.50</u>
y_2	(0.49,**0.5**,0.51), <u>0.50</u>

Table 4 Fuzzy conditional probability $P_f(Z \mid X, Y)$

X	Y	$P_f(Z = z_1 \mid X, Y)$	$P_f(Z = z_2 \mid X, Y)$
x_1	y_1	(0.64,**0.65**,0.68), <u>0.655</u>	(0.32,**0.35**,0.36), <u>0.345</u>
	y_2	(0.19,**0.2**,0.23), <u>0.205</u>	(0.77,**0.8**,0.81), <u>0.795</u>
x_2	y_1	(0.19,**0.2**,0.23), <u>0.205</u>	(0.77,**0.8**,0.81), <u>0.795</u>
	y_2	(0.04,**0.05**,0.08), <u>0.055</u>	(0.92,**0.95**,0.96), <u>0.945</u>

In Table 5, V=v_1 means that "Support Vessel (*V*)" malfunctions are observed, which may cause the support vessel collision with the FPSO unit. A triangular fuzzy number (0.34, 0.35, 0.38) is used to represent the linguistic term "*Unlikely*" which is given by experts to describe the occurrence likelihood of "Support Vessel (*V*)" malfunctions.

Table 5 Fuzzy prior probability $P_f(V)$

V	$P_f(V)$
v_1	(0.34,**0.35**,0.38), <u>0.355</u>
v_2	(0.62,**0.65**,0.66), <u>0.645</u>

The conditional fuzzy probabilities associated with the node "Collision with FPSO (*U*)" given the conditions of states of nodes "Shuttle Tanker Position (*Z*)" and "Malfunction of Support Vessel (*V*)" are shown in Table 6. There are eight fuzzy probabilities for each possible combination of states of nodes *Z* and *V*. The collision accident of the FPSO is "*Nearly Impossible*" (with fuzzy probability (0.04, 0.05, 0.08)) to happen given $Z=z_2$ and $V=v_2$ (shown in the fifth row and third column of Table 6). However once there are positioning errors of the shuttle tanker ($Z=z_1$) and the support vessel ($V=v_1$), the occurrence likelihood of the collision between the FPSO and them is increased to the fuzzy probability (0.24, 0.25, 0.28) (shown in the second row and third column of Table 6).

Table 6 Fuzzy conditional probability $P_f(U|Z,V)$

Z	V	$P_f(U=u_1\|Z,V)$	$P_f(U=u_2\|Z,V)$
z_1	v_1	(0.24,**0.25**,0.28), *0.255*	(0.72,**0.75**,0.76), *0.745*
	v_2	(0.19,**0.2**,0.23), *0.205*	(0.77,**0.8**,0.81), *0.795*
z_2	v_1	(0.19,**0.2**,0.23), 0.205	(0.77,**0.8**,0.81), *0.795*
	v_2	(0.04,**0.05**,0.08), *0.055*	(0.92,**0.95**,0.96), *0.945*

The two possible values (w_1 or w_2) of variable "Personnel Injury/Loss (*W*)" are shown in Table 7. As can be seen in the second row and second column, the occurrence likelihood of node *W* ("Personnel Injury/Loss") is "*Quite Likely*" with fuzzy probability (0.74, 0.75, 0.78) given $U= u_1$.

Table 7 Fuzzy conditional probability $P_f(W|U)$

U	$P_f(W=w_1\|U)$	$P_f(W=w_2\|U)$
u_1	(0.74,**0.75**,0.78), *0.755*	(0.22,**0.25**,0.26), *0.245*
u_2	(0.001,**0.002**,0.005), *0.0025*	(0.995,**0.998**,0.999), *0.9975*

6.4. Transformation of fuzzy numbers into crisp values

Transformation functions provided in Section 4 are used in this case study. For demonstration purposes, fuzzy prior probabilities and CPTs given in Table 2 – Table 7 are transformed into crisp numbers using Equation (8) or Equation (9), respectively.

When f-weighted valuation function takes the form of $f(\alpha)=\alpha^q$ (q=0), Equation (8) must be used. Fuzzy probability $P_f(X=x_1)=(0.39,0.4,0.43)$, for example, is transformed as follows:

$$P(X=x_1)=Val(F(x,(a,b,c,d)))$$

$$=\frac{\left(\frac{0.4+0.4}{2}\right)+\left(\frac{0.39+0.43}{2}\right)}{2}=0.405$$

The transformed prior probabilities of nodes *X*, *Y* and *V* are shown in italic and underlined fonts in Tables 2, 3 and 5, respectively. The transformed conditional probabilities of nodes *Z*, *U* and *W* are represented by italic and underlined fonts in Tables 4, 6 and 7, respectively.

When parameter q is increased to infinite (i.e. f-weighted valuation function $f(\alpha) = \alpha^q$ ($q = +\infty$)), Equation (9) must be used. Fuzzy probability $P_f(X = x_1) = (0.39, 0.4, 0.43)$ is then transformed as follows:

$$P(X = x_1) = Val(F(x,(a,b,c,d))) = \frac{0.4 + 0.4}{2} = 0.4$$

The transformed prior probabilities of nodes *X*, *Y* and *V* are shown in bold fonts in Tables 2, 3 and 5, respectively. The transformed conditional probabilities of nodes *Z*, *U* and *W* are represented by bold fonts in Tables 4, 6 and 7, respectively.

6.5. Bayesian inference

Having obtained the above probabilities, Bayesian inference rules can now be used to conduct various types of analysis. The most important use of BN is in revising probabilities in the light of actual observations of events. It is therefore possible to determine, for example, the posterior probability of human errors when it is known that there is human injury/loss observed.

The starting point of the inference is to calculate marginal probabilities. The marginal probabilities of the variables in the BN can be computed using Equation (3).

For node *Z*:

$$P(Z = z_1) = \sum_{X,Y} P(X;Y;Z = z_1) = P(X = x_1; Y = y_1; Z = z_1) + P(X = x_2; Y = y_1; Z = z_1)$$

$$+ P(X = x_1; Y = y_2; Z = z_1) + P(X = x_2; Y = y_2; Z = z_1)$$

It is worth noting that the joint probabilities of the nodes in BN structure are calculated through the conditional independence rule of probability using Equation (1):

$$P(X;Y;Z) = P(X) \times P(Y|X) \times P(Z|X;Y)$$

Considering the independent relationship between node *X* and node *Y*, the following is obtained:

$$P(X;Y;Z) = P(X) \times P(Y) \times P(Z|X;Y)$$

Hence:

$$P(Z = z_1) = P(X = x_1) \times P(Y = y_1) \times P(Z = z_1 | X = x_1; Y = y_1)$$

$$+ P(X = x_2) \times P(Y = y_1) \times P(Z = z_1 | X = x_2; Y = y_1)$$

$$+ P\ (X = x_1) \times P\ (Y = y_2) \times P\ (Z = z_1 \mid X = x_1; Y = y_2)$$

$$+ P\ (X = x_2) \times P\ (Y = y_2) \times P\ (Z = z_1 \mid X = x_2; Y = y_2)$$

For demonstration purposes, the calculation of the first part of the above equation is addressed here in detail. The following transformed probability values can be obtained from those in bold fonts of Tables 2, 3 and 4.

$P\ (X = x_1) = 0.40$; $P\ (Y = y_1) = 0.50$; $P\ (Z = z_1 \mid X = x_1; Y = y_1) = 0.65$

Therefore:

$$P\ (X = x_1) \times P\ (Y = y_1) \times P\ (Z = z_1 \mid X = x_1; Y = y_1) = 0.13$$

Then using Equation (1), $P\ (Z = z_1)$ can be calculated as:

$$P\ (Z = z_1) = 0.40 \times 0.50 \times 0.65 + 0.60 \times 0.50 \times 0.20 + 0.40 \times 0.50 \times 0.20 +$$

$$0.60 \times 0.50 \times 0.05 = 0.2450$$

Therefore,

$$P\ (Z = z_2) = 1 - P\ (Z = z_1) = 0.7550$$

For node *U*:

$$P\ (U = u_1) = \sum_{Z,V} P\ (Z; V; U = u_1) = P\ (Z = z_1; V = v_1; U = u_1) + P\ (Z = z_2; V = v_1; U = u_1)$$

$$+ P\ (Z = z_1; V = v_2; U = u_1) + P\ (Z = z_2; V = v_2; U = u_1)$$

$= 0.1307$

$$P\ (U = u_2) = 1 - P(U = u_1) = 0.8693$$

For node *W*:

$$P\ (W = w_1) = \sum_{U} P\ (U; W = w_1) = P\ (U = u_1; W = w_1) + P\ (U = u_2; W = w_1)$$

$$= P\ (U = u_1) \times P\ (W = w_1 \mid U = u_1) \times P\ (U = u_2) \times P\ (W = w_1 \mid U = u_2) = 0.0997$$

$$P\ (W = w_2) = 1 - P(W = w_1) = 0.9003$$

What is really of interest, however, is how the posterior probabilities change when new observations are added into the BN for a particular node. Suppose it is observed that there is human injury/loss, and it is required to inference to what degree the observed human injury/loss was caused by human errors during operations. This needs to calculate posterior probability $P\,(X = x_1 \mid W = w_1)$ using Bayesian rule Equation (4):

$$P\,(X = x_1 \mid W = w_1) = \frac{P\,(X = x_1; W = w_1)}{P\,(W = w_1)}$$

Therefore $P\,(X = x_1; W = w_1)$ must be calculated first. In Figure 2, it can be seen that the states of nodes *X* and *W* are affected by the states of nodes *Y, Z, V* and *U*. Based on Bayes' theorem and the Equation (2):

$$P\,(X = x_1; W = w_1) = \sum_{Y,Z,V,U} P\,(w_1 \mid U) \times P\,(U \mid Z, V) \times P\,(V) \times P\,(Z \mid x_1, Y) \times P\,(x_1) \times P\,(Y)$$

= 0.0461

Thus:

$$P\,(X = x_1 \mid W = w_1) = \frac{P\,(X = x_1; W = w_1)}{P\,(W = w_1)} = \frac{0.0461}{0.0997} = 0.4624$$

Comparing the posterior probability $P\,(X = x_1 \mid W = w_1) = 0.4624$ with prior probability $P\,(X = x_1) = 0.4$, it can be seen that there is an increase in the occurrence likelihood of human errors when a personnel injury/loss accident has been observed. This might imply that node "Personnel Injury/Loss" is quite sensitive to node "Human Errors". It can be further explained as that human errors have an impact on personnel injury/loss, although the two nodes are not linked directly in this case study. Once a personnel injury/loss accident caused by the FPSO collision took place, it is more likely that human errors are the direct/indirect cause.

It must be noted that the crisp values used in the above analysis are obtained by using Equation (9). For comparing purposes, Equation (8) is then used to do fuzzy-to-crisp values transformation. Using data in the italic and underlined fonts provided in Table 2 – Table 7, Bayesian inference can be conducted in a similar way to the previous analysis. For example:

$$P\,(X = x_1 \mid W = w_1) = \frac{P\,(X = x_1; W = w_1)}{P\,(W = w_1)} = 0.4640$$

Comparing the posterior probability $P\,(X = x_1 \mid W = w_1) = 0.4640$ with prior probability $P\,(X = x_1) = 0.405,$ it can be seen that there is an increase in the occurrence likelihood of human errors when a personnel injury/loss accident has been observed. It may suggest that there is a strong correlation between the nodes "Personnel Injury/Loss" and "Human Errors". This result is consistent to the one obtained by using Equation (9).

Having been discussed in Section 4, Equation (8)/(9) is a special case of Equation (7) when $q = 0$ or $q = +\infty$ in f-weighted valuation function ($f(\alpha) = \alpha^q$). Further calculation shows that if q takes the value between 0 to $+\infty$, the posterior probability $P(X = x_1 | W = w_1)$ takes the value between 0.4624 and 0.4640. This shows that there are no dramatic changes in analysis results when q takes different values. This demonstrates that the form of $f(\alpha) = \alpha^q$ is reasonable.

To further validate the proposed method, sensitivity analysis will be conducted in the next section.

6.6. Sensitivity analysis

Sensitivity analysis refers to a procedure to determine the sensitivity of the outcomes of an alternative to changes in its parameters. If a small change in a parameter results in relatively large changes in the outcomes, the outcomes are said to be sensitive to that parameter. This may mean that the parameter has to be determined very accurately or that the alternative has to be redesigned for low sensitivity. Sensitivity analysis is particularly useful in investigating the effects of inaccuracies or incompleteness in the parameters of a BN model on the model's output.

The most natural way of conducting sensitivity analysis is to change the parameters' values and then monitor the effects of these changes on the posterior probabilities. In this case study, the preliminary conclusion (i.e. the node "Personnel Injury/Loss" is quite sensitive to the node "Human Errors") is drawn based on posterior probabilities e.g. $P(X = x_1 | W = w_1)$. Thus one of the most important sensitivity analysis aspects is to analyse how the posterior probabilities change when the prior probabilities take different values. Without loss of generality, the fuzzy number $P_f(X = x_1)$ takes eight different values, ranging from (0.14, 0.15, 0.18) to (0.49, 0.50, 0.53) (shown in Table 8). Using the proposed method and Equations (1) – (4), the BN inference can be conducted. The numerical results are shown in Table 8 and the graphic demonstrations are shown in Figure 3 and Figure 4. As can be seen in Table 8, the pair-wise values (in the third and fourth columns, the fifth and sixth columns, respectively) clearly indicate that the posterior probability ($P(X = x_1 | W = w_1)$) steadily increases with $P(X = x_1)$. This trend can be seen in Figure 3 and Figure 4.

7. Conclusions

A novel offshore risk analysis approach is presented in this paper. The risk caused by hazardous events is modelled using BN. This approach concentrates on establishing the causal relationships among the main hazards threatening the offshore systems. It provides a mean for screening the safety implications, which would influence the development of the early design concept and normal operations. This will be suitable for carrying out offshore risk analysis associated with incomplete safety information in the early design stages or a system with a high level of fuzzy and random information. BN is a technology with huge potential for application across many domains. The proposed approach uses linguistic terms and fuzzy number-based probabilities to conduct Bayesian inference. It includes risk factor inter-relationship identification, BN model establishment, fuzzy-to-crisp probability value transformation, and Bayesian inference and interpretation. The BN model explicitly represents the cause-and-effect relationships among system variables that may be obscured under other modelling approaches such as rule-based fuzzy reasoning. The flexibility of the approach allows for multiple forms of information to be used to quantify model relationships, including formally assessed expert opinion when quantitative data are lacking or, only qualitative/vague statements can be made. Particularly, the proposed f-weighted valuation function takes into account not only the dominating values, but also the $\alpha - level - values$ that are ignored by conventional evaluation methods.

Table 8 Sensitivity analysis results between $P\ (X = x_1 \mid W = w_1)$ and $P_f(X = x_1)$

	Fuzzy prior probabilities	Crisp prior probabilities and posterior probabilities under *q=0*		Crisp prior probabilities and posterior probabilities under $q = +\infty$	
No	$P_f(X = x_1)$	$P(X = x_1)$	$P(X = x_1 \mid W = w_1)$	$P(X = x_1)$	$P(X = x_1 \mid W = w_1)$
1	(0.14,0.15,0.18)	0.15	0.1852	0.155	0.1891
2	(0.19,0.20,0.23)	0.20	0.2493	0.205	0.2469
3	(0.24,0.25,0.28)	0.25	0.3004	0.255	0.3033
4	(0.29,0.30,0.33)	0.30	0.3529	0.305	0.3582
5	(0.34,0.35,0.38)	0.35	0.4093	0.355	0.4117
6	(0.39,0.40,0.43)	0.40	0.4624	0.405	0.4640
7	(0.44,0.45,0.48)	0.45	0.5091	0.455	0.5149
8	(0.49,0.50,0.53)	0.50	0.5634	0.505	0.5647

The proposed BN model is also a modular representation of uncertainty knowledge caused due to randomness and vagueness, and therefore can be potentially integrated into the existing risk assessment systems. In general, the proposed model ensures that its applications are conducted in a disciplined, well-managed, and consistent manner that promotes the delivery of risk assessment results.

Acknowledgements
This project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant References: GR/S85504 and GR/S85498.

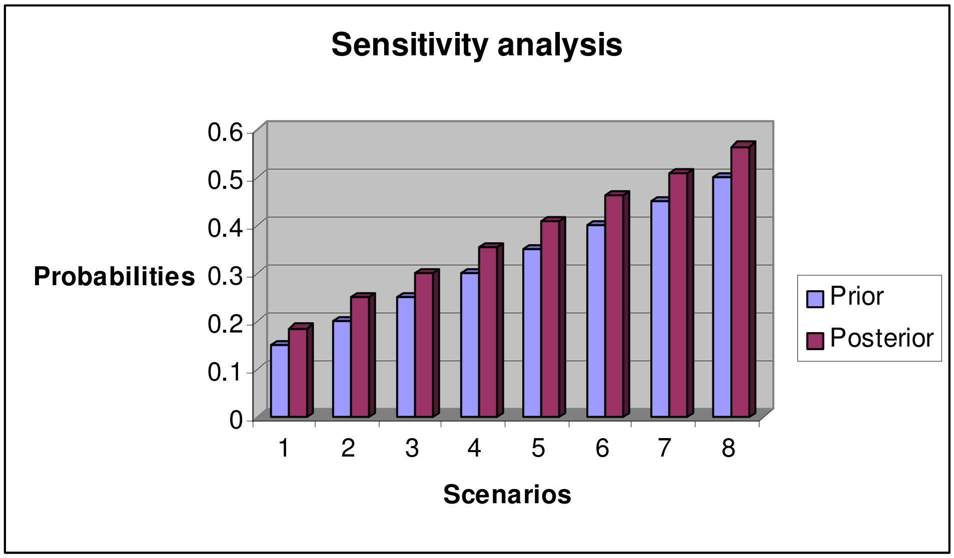

Figure 3: Sensitivity analysis between $P(X = x_1 | W = w_1)$ and $P(X = x_1)$ under $q = 0$

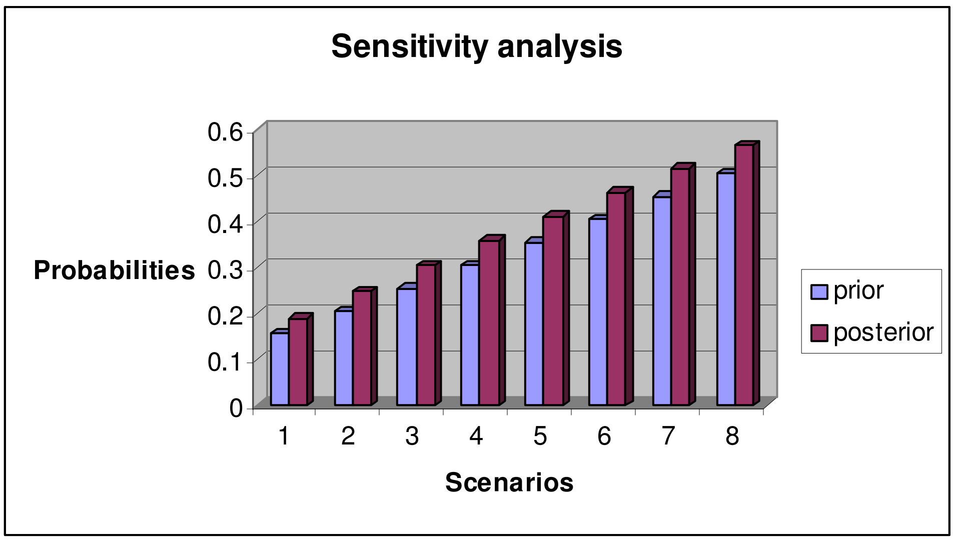

Figure 4: Sensitivity analysis between $P(X = x_1 | W = w_1)$ and $P(X = x_1)$ under $q = +\infty$

References

Detynieckim, M., and Yager, R. (2000). "Ranking fuzzy numbers using alpha-weighted valuations." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(5), 573-592.

Hayes, K. R. (1998). "Bayesian statistical inference in ecological risk assessment." Crimp Technical Report (No.17), Centre for Research on Introduced Marine Pests, CSIRO, Hobart, Australia.

Heckerman, D. (1997). "Bayesian networks for data mining." *Data Mining and Knowledge Discovery*, 1(1), 79-119.

Heckerman, D., and Breese, J. (1996). "Causal independence for probability assessment and inference using Bayesian networks." *IEEE Transactions on Systems, Man, and Cybernetics*, 26(6), 826-831.

HSE. (1996). *The offshore installations and wells (design and construction, etc.) regulations 1996*, Health &Safety Executive, HSE Books, ISBN: 011054451X, No. 913, London.

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

Kadie, C. M., Hovel, D., and Horvitz, E. (2001). "MSBNx: a component-centric toolkit for modelling and inference with Bayesian networks." *MSR-TR-2001-67*, Microsoft Corporation, Redmond, WA.

Kang, C. W., and Golay, M. W. (1999). "A Bayesian belief network - based advisory system for operational availability focused diagnosis of complex nuclear power systems." *Expert Systems with Applications*, 17(1), 21-32.

Karwowski, W., and Mital, A. (1986). "Potential applications of fuzzy sets in industrial safety engineering." *Fuzzy Sets and Systems*, 19(2), 105-120.

Lauritzen, S. L., and Spiegelhalter, D. J. (1988). "Local computations with probabilities on graphical structures and their application to expert systems." *Journal of the Royal Statistical Society, Series B*, 50(2), 157-224.

Mitchell, K., and Bright, C. K. (1995). "Minimizing the potential for human error in ship operation." *IMAS 95, Management and Operation of Ships, Practical Techniques for Today and Tomorrow*.

Myllymaki, P. (2005). "Advantages of Bayesian networks in data mining and knowledge discovery." http://www.bayesit.com/docs/advantages.html, (access date: 11/06/2006).

Netica. (2002). "Netica-J reference manual - version 2.21." Java Version of Netica API, Norsys Software Corporation.

Pearl, J. (1988). "Probabilistic reasoning in intelligent systems." *Networks of Plausible Inference*, Morgan Kaufmann, Palo Alto, USA.

Ren, J., Jenkinson, I., Sii, H. S., Xu, D. L., Wang, J., and Yang, J. B. (2005a). "An offshore safety assessment framework using fuzzy reasoning and evidential synthesis approaches." *Journal of Marine Engineering and Technology (Part A)*(A6), 3-16.

Ren, J., Wang, J., and Jenkinson, I. (2005b). "Fuzzy Bayesian modelling in maritime risk analysis." *GERI Annual Research Symposium*, Liverpool John Moores University, UK.

Ross, T. J. (1995). *Fuzzy logic with engineering applications*, McGraw-Hill, Inc.

Sii, H. S., Ruxton, T., and Wang, J. (2002). "A synthesis using fuzzy set theory and Dempster-Shafer based approach to compromise decision making with multiple attribute applied to risk control options selection." *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 216(1), 15-29.

UKDE. (1990). "The public inquiry into the Piper Alpha disaster (Cullen Report)." *ISBN 010113102*, UK Department of Energy, London, UK.

UKOOA. (1999). "Industry guidelines on a framework for risk related decision making." UK Offshore Operators Association, Aberdeen, UK.

Wang, J. (2005). "Maritime risk assessment and its current states." *Quality and Reliability Engineering International*, 21(7), 1-17.

Wang, J., and Kieran, O. (2000). "Offshore safety assessment and safety based decision making - the current status and future aspects." *Journal of Offshore Mechanics and Arctic Engineering - Transactions of the ASME*, 122(2), 93-99.

Wang, J., and Ruxton, T. (1997). "Design for safety." *Journal of American Society of Safety Engineers*, Jan., 23-29.

Wang, J., Yang, J. B., and Sen, P. (1995). "Safety analysis and synthesis using fuzzy set modelling and evidential reasoning." *Reliability Engineering and System Safety*, 47(2), 103-118.

Yager, R. R. (1981). "A procedure for ordering fuzzy subsets of the unit interval." *Information Sciences*, 24(2), 143-161.

Yager, R. R. (1999). "On ranking fuzzy numbers using valuations." *International Journal of Intelligent Systems*, 14(6), 1249-1268.

Yang, J. B., and Xu, D. L. (2002). "On the evidential reasoning algorithm for multiattribute decision analysis under uncertainty." *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 32(3), 289-304.

Zhang, G., Bai, C., Lu, J., and Zhang, C. (2004). "Bayesian network based cost benefit factor inference in e-services." *Proceedings of the 2nd International Conference on Information Technology for Application (ICITA)*, Harbin, P. R. China, January 7-10, 464-469.