Ann Oper Res (2008) 163: 143–168
DOI 10.1007/s10479-008-0327-0

Self-tuning of fuzzy belief rule bases for engineering system safety analysis

Jun Liu · Jian-Bo Yang · Da Ruan · Luis Martinez · Jin Wang

Published online: 11 March 2008

Abstract A framework for modelling the safety of an engineering system using a fuzzy rule-based evidential reasoning (*FURBER*) approach has been recently proposed, where a fuzzy rule-base designed on the basis of a belief structure (called *a belief rule base*) forms a basis in the inference mechanism of *FURBER*. However, it is difficult to accurately determine the parameters of a fuzzy belief rule base (FBRB) entirely subjectively, in particular for complex systems. As such, there is a need to develop a supporting mechanism that can be used to train in a locally optimal way a FBRB initially built using expert knowledge. In this paper, the methods for self-tuning a FBRB for engineering system safety analysis are investigated on the basis of a previous study. The method consists of a number of *single* and *multiple objective nonlinear optimization* models. The above framework is applied to model the system safety of a marine engineering system and the case study is used to demonstrate how the methods can be implemented.

Keywords Safety analysis · Uncertainty · Fuzzy logic · Belief rule-base · Evidential reasoning · Optimization

J. Liu (✉)
School of Computing and Mathematics, Faculty of Computing and Engineering, University of Ulster at Jordanstown, Newtownabbey BT37 0QB, Northern Ireland, UK
e-mail: j.liu@ulster.ac.uk

J.-B. Yang
Manchester Business School (East), The University of Manchester, Manchester M15 6PB, UK

D. Ruan
Belgian Nuclear Research Centre (SCK•CEN), Boeretang 200, 2400 Mol, Belgium

L. Martinez
Department of Computer Science, University of Jaén, 23071 Jaén, Spain

J. Wang
School of Engineering, Liverpool John Moores University, Liverpool, UK

1 Introduction

The growing technical complexity of large engineering systems such as offshore platforms and offshore support vessels, together with the intense public concern over their safety has stimulated the research and development of novel safety analysis methods and safety assessment procedures.

The safety of a large engineering system is affected by many factors regarding its design, manufacturing, installation, commissioning, operation and maintenance. Consequently, it may be extremely difficult to construct an accurate and complete mathematical model for the system in order to assess the safety because of inadequate knowledge about the basic failure evens. This leads inevitably to problems of uncertainty in representation (Wang et al. 1995).

In engineering safety analysis, intrinsically vague information may coexist with conditions of "lack of specificity" originating from evidence not strong enough to completely support a hypothesis but only with degrees of belief or *credibility* (Binaghi and Madella 1999; Liu et al. 2003). One realistic way to deal with imprecision is to use linguistic assessments instead of numerical values. Fuzzy logic approaches (Zadeh 1965, 1975a, 1975b, 1975c) employing fuzzy IF-THEN rules can model the qualitative aspects of human knowledge and reasoning process without employing precise quantitative analysis. This actually provides a tool for working directly with the linguistic information, which is commonly used for representing risk factors and carrying out safety assessments (Bell and Badiru 1996; Bowles and Pelaez 1995; Duckstein 1994; Wang et al. 1995). Dempster-Shafer (D-S) theory of evidence (Dempster 1968; Shafer 1976) based on the concept of *belief function* is well suited to modeling subjective credibility induced by partial evidence (Smets 1988). It also provides appropriate methods for computing belief functions for combination of evidence. Besides, the D-S theory also shows great potentials in multiple attribute decision analysis (MADA) under uncertainty, where the evidential reasoning (ER) approach for MADA under uncertainty has been developed, on the basis of a distributed assessment framework and the evidence combination rule of the D-S theory (Yang and Singh 1994; Yang and Sen 1994; Yang 2001; Yang and Xu 2002a, 2002b; Yang et al. 2006b).

Accordingly, it seems reasonable to extend the fuzzy logic framework to cover credibility uncertainty as well. In order to combine fuzzy logic and D-S models to deal with fuzziness and incompleteness in safety analysis, a framework for modelling the safety of an engineering system using a fuzzy rule-based evidential reasoning (*FURBER*) approach was recently proposed (Liu et al. 2004), which is based on a generic Rule-base Inference Methodology using the Evidential Reasoning approach (*RIMER*) developed recently in Yang et al. (2006a). In the framework, a fuzzy rule-base designed on the basis of a belief structure, called *a belief rule base*, is used to capture uncertainty and non-linear relationships between the parameters, and the inference of the rule-based system is implemented using the evidential reasoning algorithm proposed in Yang (2001), Yang and Xu (2002a, 2002b).

A belief rule base forms a basis in the inference mechanism of *FURBER*, which is a framework for representing expert knowledge. However, it is difficult to determine its elements entirely subjectively, in particular for a large scale fuzzy rule base with hundreds of rules. Also, a change in a rule weight or an attribute weight may lead to significant changes in the performance of a belief rule base. Moreover, the forms of fuzzy membership functions in the antecedent of a rule still remain an important factor for the system performance.

As such, there is a need to develop a method that can be used to train a FBRB in an optimal way using expert judgments as well as statistical data, especially for safety analysis of a large complex engineering system. Recently, optimal models for training the elements of general belief rule bases and other knowledge representation parameters in *RIMER* have been proposed in Yang et al. (2006c). Based on these optimal models, revised models for self-tuning a FBRB in *FURBER* for engineering system safety analysis are investigated in this paper. In a FBRB, input data, attribute weights, rule weights, and parameters of fuzzy membership functions are combined to generate activation weights for rules, and all activated belief rules are then combined to generate appropriate conclusions using the evidential reasoning approach. Such a combination process is formulated as nonlinear objective functions to minimize the differences between observed outputs and the outputs of a belief rule base whilst parameter specific limits and partial expert judgments can be formulated as constraints. The optimization problems can be solved using existing tools such as the optimization tool box provided in *MATLAB*.

The rest of this paper is organized as follows. A safety analysis framework and *FURBER* is briefly reviewed in Sect. 2. The optimization method for constructing a FBRB is investigated in Sect. 3. A numerical example is illustrated in Sect. 4. Some discussions are given in Sect. 5. Conclusions are drawn in Sect. 6.

2 Safety analysis framework and *FURBER*

This section reviews the *FURBER* framework (Liu et al. 2004). The proposed framework for modelling system safety consists of four major components, which outline all the necessary steps required for safety evaluation using fuzzy rule-based evidential reasoning approach.

2.1 Identify causes/factors

This can be done by a panel of experts during a brainstorming session at the early concept design stages of the system.

2.2 Identify & definite fuzzy input and fuzzy output variables (i.e., safety estimates)

The three fundamental parameters used to assess the safety level of an engineering system on a subjective basis are the *failure rate* (**FR**)*, consequence severity* (**CS**) and *failure consequence probability* (**FCP**). Subjective assessments (using linguistic variables instead of ultimate numbers in probabilistic terms) are more appropriate for safety analysis as they are always associated with great uncertainty, especially for a novel system with high level of innovation (Bell and Badiru 1996; Bowles and Pelaez 1995; Duckstein 1994; Wang et al. 1995; Sii and Wang 2002; Liu et al. 2003). These linguistic assessments can become the criteria for measuring safety levels. The typical linguistic variables used to describe **FR**, **CS**, and **FCP** of a particular element may be described as follows respectively:

FR describes failure frequencies in a certain period, which directly represents the number of failures anticipated during the design life span of a particular system or an item. To estimate the **FR**, one may choose to use such linguistic terms as "*very low,*" "*low,*" "*reasonably low,*" "a*verage,*" "*reasonably frequent,*" "*frequent,*" and "*highly frequent.*"

CS describes the magnitude of possible consequences, which is ranked according to the severity of failure effects. One may choose to use such linguistic terms as "*negligible,*" "*marginal,*" "*moderate,*" "*critical,*" and "*catastrophic.*"

FCP defines the probability that consequences happen given the occurrence of the event. One may choose to use such linguistic terms as "*highly unlikely,*" "*unlikely,*" "*reasonably unlikely,*" "*likely,*" "*reasonably likely,*" and "*definite.*"

The detailed definitions of these parameters and their membership functions can be found in Sii and Wang (2002), Liu et al. (2004).

Safety estimate is the only output fuzzy variable used in this study to produce safety evaluation for a particular cause to technical failure. This variable is also described linguistically, which is described and determined by the above parameters. In safety assessment, it is common to express a safety level by degrees to which it belongs to such linguistic variables as "*poor,*" "*fair,*" "*average,*" and "*good*" that are referred to as safety expressions.

2.3 Fuzzy rule-base with the belief structure

Fuzzy logic systems are knowledge-based or rule-based ones constructed from human knowledge in the form of fuzzy *IF-THEN* rules. For example, the following is a simple fuzzy *IF-THEN* rule for safety estimate:

IF **FR** of a hazard is *frequent* AND **CS** is *catastrophic* AND **FCP** is *likely,*
THEN **safety estimate** is *Poor.*

In view of the increasing complexity of many knowledge-based systems, the knowledge representation power of fuzzy rule-based systems will be severely limited if only fuzziness is used to represent uncertain knowledge. As mentioned in the first section, there is another kind of uncertainty caused because an expert is unable to establish a strong correlation between premise and conclusion. In other words, evidence available is not sufficient or experts are not 100% certain to believe in a hypothesis but only to degrees of belief. For example, we may only get the following rule with certain degrees of belief:

R_1: IF the **FR** is *frequent* AND the **CS** is *critical* AND the **FCP** is *unlikely*
THEN **safety estimate** is *Fair* with a belief degree of 0.7.

More generally, we may have fuzzy rules with belief degrees for multiple possible consequent terms, for example,

R_k: IF the **FR** is *frequent* AND the **CS** is *critical* AND the **FCP** is *unlikely*
THEN **safety estimate** is $\{(Good, 0), (Average, 0), (Fair, 0.7), (Poor, 0.3)\}$

where $\{(Good, 0), (Average, 0), (Fair, 0.7), (Poor, 0.3)\}$ is a belief distribution representation for safety consequent, representing that we are 70% sure that safety level is *Fair*, and 30% sure that safety level is *Poor*.

In order to model more general and complex decision-making problems under uncertainty, other important information such as weighting factors may also need to be considered, such as the relative weight of a rule (rule weight) used to represent the relative importance of the rule's contribution to reach the final conclusion, and the relative weight of an antecedent attribute (attribute weight).

To take into account the belief degrees of a rule, attribute weights and rule weights, a simple fuzzy rule for safety estimate is extended to a so-called *fuzzy belief rule* with all possible consequents associated with belief degrees. In general, assume that the three antecedent parameters, $U_1 = \mathbf{FR}$, $U_2 = \mathbf{CS}$ and $U_3 = \mathbf{FCP}$, can be described by J_i linguistic terms $\{A_{ij}; j = 1, \ldots, J_i\}$, $i = 1, 2, 3$, respectively. One consequent variable ***safety estimates*** can be described by N linguistic terms, i.e., $D_1, D_2, \ldots, D_N$. Suppose that the

Table 1 A belief rule expression matrix

Belief output	Input					
	$\boldsymbol{A^1}(w_1)$	$\boldsymbol{A^2}(w_2)$	$\cdots$	$\boldsymbol{A^k}(w_k)$	$\cdots$	$\boldsymbol{A^L}(w_L)$
D_1	β_{11}	β_{12}	$\cdots$	β_{1k}	$\cdots$	β_{1L}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
D_i	β_{i1}	β_{i2}	$\cdots$	β_{ik}	$\cdots$	β_{iL}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$
D_N	β_{N1}	β_{N2}	$\cdots$	β_{Nk}	$\cdots$	β_{NL}

rule-base is given by $R = \{R_1, R_2, \ldots, R_L\}$, the kth rule can be represented in a compact format as follows:

$$R_k: \quad \text{IF } \boldsymbol{U} \text{ is } \boldsymbol{A^k} \text{ THEN } \textbf{safety estimates} \text{ is } \boldsymbol{D} \text{ with belief degree} \beta^k \text{, with a rule weight } \theta_k \text{ and attribute weights } \delta_1, \delta_2, \delta_3 \tag{1}$$

where $\boldsymbol{U}$ represents the antecedent attribute vector (U_1, U_2, U_3), $\boldsymbol{A^k}$ the packet antecedents $\{A_1^k, A_2^k, A_3^k\}$, where A_i^k $(\in \{A_{ij}; j = 1, \ldots, J_i\})$ is a linguistic term corresponding to the ith attribute in the kth rule with $i = 1, 2, 3$. $\boldsymbol{D}$ is the consequent vector $(D_1, \ldots, D_N)$, and β^k the vector of the belief degrees $(\beta_{1k}, \ldots, \beta_{Nk})$ for $k \in \{1, \ldots, L\}$ with$\sum_{i=1}^{N} \beta_{ik} \le 1$. This is the vector form of *a belief rule*, β_{ik} measures the degree to which D_i is the consequent if the input activates the packet antecedent $\boldsymbol{A^k}$ in the kth rule for $i = 1, \ldots, N; k = 1, \ldots, L$. L is the number of rules in the rule-base. If $\sum_{i=1}^{N} \beta_{ik} = 1$, the output assessment or the kth rule is said to be complete; otherwise, it is incomplete. The fuzzy rule-base in the form (1) is referred to as *a fuzzy belief rule-base* where either $\boldsymbol{U}$ or $\boldsymbol{D}$ or both could be fuzzy.

A fuzzy belief rule base given in the form shown in (1) represents functional mappings between antecedents and consequents with uncertainty. It provides a more informative and realistic scheme for uncertain knowledge representation. Note that the degrees of belief β_{ik} $(i = 1, \ldots, N; k = 1, \ldots, L)$, and the weights could be assigned initially by experts and then trained or updated using dedicated learning algorithms if system input and output information is available (Yang et al. 2006c). It is the purpose of this paper to investigate such optimal learning models for training a FBRB. Once a FBRB is constructed and trained, its knowledge contained in the FBRB can be used to perform inference for given inputs.

The packet antecedent A^k of a belief *IF-THEN* rule can be considered as a global attribute, which is considered to be assessed to a consequent D_i with a belief degree of $\beta_{ik}(i \in \{1, \ldots, N\})$. This assessment can be represented by

$$S(\boldsymbol{A^k}) = \{(D_i, \beta_{ik}); i = 1, \ldots, N\}, \tag{2}$$

which is a distribution assessment and is referred to as *a fuzzy belief structure* if D_i is a fuzzy set, where β_{ik} measures the degree to which D_i is the consequent if the input activates the antecedent $\boldsymbol{A^k}$ in the kth rule for $i = 1, \ldots, N$, $k = 1, \ldots, L$. L is the number of rules in the rule-base and N is the number of possible consequents.

Suppose that all the L rules are independent of each other, which means that the packet antecedent $\boldsymbol{A^1}, \ldots, \boldsymbol{A^L}$ are independent of each other. The fuzzy belief rule-base given by (1) and (2) can then be summarized using a fuzzy belief rule expression matrix as shown in Table 1, where w_k is the activation weight of $\boldsymbol{A^k}$, which is similar to the belief rule expression matrix (Yang et al. 2006a, 2006c) and measures the degree to which the kth rule is weighted and activated.

2.4 Fuzzy rule-base inference mechanism based on the evidential reasoning approach

In an established FBRB, input of an antecedent is transformed into a distribution on the linguistic values of this antecedent. Such a distribution describes the degree of the antecedent being activated. Moreover, the antecedents of an *IF-THEN* rule form an overall attribute, i.e., *a packet antecedent attribute*. The activation weight of a rule can be generated by aggregating the degrees to which all antecedents in the rule are activated. In this context, an *IF-THEN* rule can be considered as an evaluation problem of a packet antecedent attribute being assessed to an output term in the consequent of the rule with certain degrees of belief. Finally, the inference of a rule-based system is implemented using the evidential reasoning approach (Yang et al. 2006a). It is summarized as follow.

In the belief rule expression matrix, the degree of activation of the kth rule w_k is calculated by Yang et al. (2006a):

$$w_k = (\theta_k \cdot \alpha_k) \Big/ \left(\sum_{j=1}^{L} \theta_j \alpha_j \right), \tag{3}$$

where θ_k $(\in \mathbf{R}^+, k = 1, \ldots, L)$ is the relative weight of the kth rule, $\alpha_k = \prod_{i=1}^{3} (\alpha_i^k)^{\bar{\delta}_i}$ is called *the normalized combined matching degree*, here $\bar{\delta}_i = \delta_i / (\max_{i=1,2,3}\{\delta_i\})$, so $\bar{\delta}_i \in [0, 1]$, here δ_i $(\in \mathbf{R}^+, i = 1, 2, 3)$ is the weight of the ith antecedent attribute. Note that "AND" connective is used for three antecedents in the rule. In other words, the consequent of a rule is not believed to be true unless all the antecedents of the rule are activated. In such cases, the simple multiplicative aggregation function is used here to calculate α_k.

Here θ_k and δ_i can be assigned to any value in $\mathbf{R}^+$ depending on a specific application because w_k will be eventually normalized so that $w_k \in [0, 1]$ by (3). Without lose of generality, however, we assume that $\theta_k \in [0, 1]$, $(k = 1, \ldots, L)$ and $\delta_i \in [0, 1]$ $(i = 1, 2, 3)$. Note that $0 \le w_k \le 1$ $(k = 1, \ldots, L)$ and $\sum_{i=1}^{L} w_i = 1$.

α_i^k could be generated using various ways depending on the nature of an antecedent attribute and the available data. Note that the linguistic values of an attribute and the types of input information are problem specific and thus their definitions depend on problems in hand. More discussions about this issue can be found in the references (Yang et al. 2006a; Liu et al. 2004).

A linguistic value of an attribute may in general be regarded as an evaluation grade. An input x_i^* for an attribute X_i can be equivalently transformed to a distribution over the linguistic values defined for the attribute using belief degrees as follows (Yang 2001):

$$S(x_i^*) = \{(A_{ij}, \alpha_{ij}), j = 1, \ldots, J_i\}, \quad i = 1, 2, 3, \tag{4}$$

where A_{ij} is the jth linguistic value of the attribute X_i, α_{ij} the degree to which the input for X_i belongs to the linguistic value A_{ij} with $\alpha_{ij} \ge 0$ and $\sum_{j=1}^{J_i} \alpha_{ij} \le 1$ $(i = 1, 2, 3)$ and J_i the number of the linguistic values used for describing the ith antecedent attribute X_i. The above distributed assessment reads that the input for the attribute X_i is assessed to the linguistic value A_{ij} with the degree of belief of α_{ij} $(j = 1, 2, \ldots, J_i$ and $i = 1, 2, 3)$. If a belief rule base has T antecedent attributes, then rules in the rule-base will be normally constructed by taking all possible combinations of the linguistic values for the T attributes. Hence α_i^k is generated corresponding to the kth rule defined as in (1), i.e.,

$$\mathbf{FR} \text{ is } (A_1^k, \alpha_1^k) \text{ AND } \mathbf{CS} \text{ is } (A_1^k, \alpha_2^k) \text{ AND } \mathbf{FCP} \text{ is } (A_1^k, \alpha_3^k), \tag{5}$$

where $A_i^k \in \{A_{ij}, j = 1, \dots, J_i\}$ and $\alpha_i^k \in \{\alpha_{ij}, j = 1, \dots, J_i\}$ is the individual matching degree to which the input belongs to the linguistic value A_i^k of the individual antecedent U_i appearing in the kth rule.

As a result, each input can be represented as a distribution on linguistic values using a belief structure. The main advantage of doing so is that precise data, random numbers and subjective judgments with uncertainty can be consistently modelled under the same framework. However, the focus of this paper is to train the belief rule-base and its parameters involved. The details on how to get the individual matching degree α_i^k is discussed in Yang (2001), Yang et al. (2006a). We consider the following form as a case study.

For a given real input vector $\boldsymbol{x} = (x_1, x_2, x_3), \alpha_i^k = \mu_{A_i^k}(x_i)$, where $\mu_{A_i^k}(\bullet)$ is the fuzzy membership function of the linguistic term A_i^k. The fuzzy membership function can be applied in different forms depending on the system. In Sii and Wang (2002), Liu et al. (2004), the straight-line membership functions are used due to its advantage of simplicity, such as the triangular membership function and trapezoidal membership function. Continuous and differentiable Gaussian function is used in this paper, i.e.,

$$\mu_{A_i^k}(x_i) = \exp\left(-\frac{1}{2}\left(\frac{x_i - c_i^k}{\sigma_i^k}\right)^2\right), \tag{6}$$

where c_i^k is the central value of the fuzzy membership function and σ_i^k is the variance at the central value.

Having represented each belief rule using (2), the evidential reasoning (ER) approach (Yang 2001;Yang and Xu 2002a, 2002b; Yang et al. 2006b) can be directly applied to combine the activated belief rules and generate final conclusions as follows. Note that in this paper the terms D_i in the consequent of a fuzzy rule is assumed to be independent, so that the following ER algorithm can be used for FBRB inference. If D_i are dependent, or there is overlap between the two adjacent fuzzy sets D_i and D_{i+1} , then the fuzzy ER algorithm (Yang et al. 2006b) should be used for FBRB inference. First, transform the degrees of belief β_{jk} for all $j = 1, \dots, N, k = 1, \dots, L$ into basic probability masses using the following ER approach (Yang 2001; Yang and Xu 2002a, 2002b):

$$m_{j,k} = w_k \beta_{j,k}, \quad j = 1, \dots, N;$$

$$m_{D,k} = 1 - \sum_{j=1}^{N} m_{j,k} = 1 - w_k \sum_{j=1}^{N} \beta_{j,k},$$

$$\bar{m}_{D,k} = 1 - w_k, \quad \text{and} \quad \tilde{m}_{D,k} = w_k\left(1 - \sum_{j=1}^{N} \beta_{j,k}\right)$$

for all $k = 1, \dots, L$, where $m_{D,k} = \bar{m}_{D,k} + \tilde{m}_{D,k}$ for all $k = 1, \dots, L$ and $\sum_j^L w_j = 1$. The probability mass assigned to the consequent set D, which is unassigned to any individual consequent, is split into two parts, one caused by the relative importance of the kth packet antecedent A^k (or $\bar{m}_{D,k}$) and the other by the incompleteness of the kth packet antecedent A^k (or $\tilde{m}_{D,k}$).

Then, aggregate all the packet antecedents of the L rules to generate the combined degree of belief in each possible consequent D_j in $\boldsymbol{D}$. Suppose $m_{j,I(k)}$ is the combined degree of belief in D_j by aggregating the first k packet antecedents $(\boldsymbol{A^1}, \ldots, \boldsymbol{A^k})$ and $m_{D,I(k)}$ is the remaining degree of belief unassigned to any consequent. Let $m_{j,I(1)} = m_{j,1}$ and $m_{D,I(1)} = m_{D,1}$. Then the overall combined degree of belief β_j in D_j is generated as follows:

$$\begin{aligned}
\{D_j\}\colon\quad & m_{j,I(k+1)} = K_{I(k+1)}[m_{j,I(k)}m_{j,k+1} + m_{j,I(k)}m_{D,k+1} + m_{D,I(k)}m_{j,k+1}],\\
& m_{D,I(k)} = \bar{m}_{D,I(k)} + \tilde{m}_{D,I(k)}, \quad k = 1, \ldots, L,\\
\{D\}\colon\quad & \tilde{m}_{D,I(k+1)} = K_{I(k+1)}[\tilde{m}_{D,I(k)}\tilde{m}_{D,k+1} + \tilde{m}_{D,I(k)}\bar{m}_{D,k+1} + \bar{m}_{D,I(k)}\tilde{m}_{D,k+1}],\\
\{D\}\colon\quad & \bar{m}_{D,I(k+1)} = K_{I(k+1)}[\bar{m}_{D,I(k)}\bar{m}_{D,k+1}],\\
& K_{I(k+1)} = \left[1 - \sum_{j=1}^{N}\sum_{\substack{t=1\\ t\neq j}}^{N} m_{j,I(k)}m_{t,k+1}\right]^{-1}, \quad k = 1, \ldots, L-1,\\
\{D_j\}\colon\quad & \beta_j = \frac{m_{j,I(L)}}{1 - \bar{m}_{D,I(L)}}, \quad j = 1, \ldots, N,\\
\{D\}\colon\quad & \beta_D = \frac{\tilde{m}_{D,I(L)}}{1 - \bar{m}_{D,I(L)}},
\end{aligned}$$

β_D represents the remaining belief degree unassigned to any D_j. It has been proved that $\sum_{j=1}^{N} \beta_j + \beta_D = 1$ (Yang and Xu 2002b). The final conclusion generated by aggregating all activated rules by the actual input vector $\boldsymbol{x} = \{x_t; t = 1, 2, 3\}$ can be represented as follows

$$\{(D_j, \beta_j), j = 1, \ldots, N\}. \tag{7}$$

The above recursive ER algorithm combines various piece of evidence on a one-by-one basis. The advantage of doing so is its clarity in concept. In situations where an explicit ER aggregation function is more desirable as in optimization model, an analytical ER algorithm will be desirable. In view of this, the ER Recursive Algorithm used in Yang and Xu (2002a, 2002b) has been equivalently transformed into the analytical ER algorithm (Wang et al. 2006). Using this analytical ER algorithm, the overall combined degree of belief β_j in D_j is generated as follows:

$$\beta_j = \frac{\mu \times [\prod_{k=1}^{L}(w_k\beta_{j,k} + 1 - w_k\sum_{i=1}^{N}\beta_{i,k}) - \prod_{k=1}^{L}(1 - w_k\sum_{i=1}^{N}\beta_{i,k})]}{1 - \mu \times [\prod_{k=1}^{L}(1 - w_k)]}, \quad j = 1, \ldots, N, \tag{8}$$

where $\mu = [\sum_{j=1}^{N}\prod_{k=1}^{L}(w_k\beta_{j,k} + 1 - w_k\sum_{i=1}^{N}\beta_{i,k}) - (N-1)\prod_{k=1}^{L}(1 - w_k\sum_{i=1}^{N}\beta_{i,k})]^{-1}$.

The analytical ER algorithm offers the ER approach more flexibility in aggregating a large number of environmental factors (or basic attributes). It clearly shows its nonlinear features (Yang and Xu 2002b) and provides a straightforward way to conduct sensitivity analysis for the parameters of the ER approach such as weights and belief degrees. It also facilitates the estimation and optimization of these parameters. However, it should be noted that in both the above recursive and analytical ER algorithms, D_j are assumed to be independent evaluation grade. If they are dependent, then the fuzzy ER algorithm (Yang et al. 2006b) should be used to generate β_j. Equation (8) looks complicated, the meanings of each parameter in the recursive ER algorithm, also in (8) can be referred to Yang (2001), Yang and Xu (2002a, 2002b). Because the inference process in the proposed method is a kind of nonlinear algorithm using ER, it will be more complicated. Hence, the assessment of the parameters of the inference procedure is not easy. That is indeed the main objective of the proposed work trying to using self-tuning to help efficiently generate near-optimal parameters.

The logic behind the approach is that if the consequent in the kth rule includes D_i with $\beta_{ik} > 0$ and the kth rule is activated then the overall output must be D_i to a certain degree. The degree is measured by both the degree to which the kth rule is important to the overall output and the degree to which the antecedents of the kth rule are activated by the actual input $\boldsymbol{x}^*$.

Note that just as RIMER was developed to enhance conventional IF-THEN rule based systems (Yang et al. 2006a), *FURBER* is developed on the basis of and to enhance conventional fuzzy IF-THEN rule based systems. Both methods are capable of modelling and simulating explicit expert knowledge using IT-THEN rules but the former can model both discrete and continuous relationships with uncertainty. The unique feature of the *FURBER* method, which differentiates it from other existing modelling methods, is that it can explicitly model and infer with ignorance (incompleteness) that may exist either inside the model structure (rules) or in input data. Ignorance can be caused due to incomplete or missing data or the inability of experts or decision makers to provide complete or accurate judgments, which is common in situations where human knowledge has to be used.

Notice that it is the beliefs used in the belief structure and the activation weights that determine the actual performance of inference. The degree to which the final output can be affected is determined by the magnitude of the activation weight and the belief degrees in each rule. On the other hand, if the parameters of a FBRB such as $\beta_{i,k}, \theta_k$ and δ_i are not given a priori or only known partially or imprecisely, they could be trained using observed input and output information. Therefore, the performance of inference can be improved if the following parameters

(1) Rule weight θ_k $(k = 1, \ldots, L)$ and attribute weights $\delta_1, \delta_2, \delta_3$;
(2) The degrees of belief β_{ik} $(i = 1, \ldots, N; k = 1, \ldots, L)$;
(3) The central value of fuzzy membership function c_{ij} and the variance σ_{ij} at the central value $(i = 1, \ldots, 3;\ j = 1, \ldots, J_i)$;

in (8) are adjusted by autonomous learning. Moreover the utility $u(D_j)$ $(j = 1, \ldots, N)$ of the linguistic term of the safety estimate can also be trained if required. This is exactly the topic for the rest of this paper.

Notice that there are some constraint conditions on each parameter in the above formulation, which are described in the optimization formulations in the following section.

3 Optimal method for belief rule bases in *FURBER*

3.1 A generic optimal framework

In this section, the optimal algorithm is to be incorporated in the *FURBER* framework, whose function is to search for optimal FBRB and other knowledge representation parameters simultaneously.

As discussed in Yang et al. (2006a, 2006c), beliefs in a belief rule base may initially be provided by human experts based on individuals' experiences and personal judgments and then optimally trained if observed input-output data is available. In other words, $\beta_{i,k}$ can be trained if appropriate data becomes available. Also, a change in rule weights θ_k and

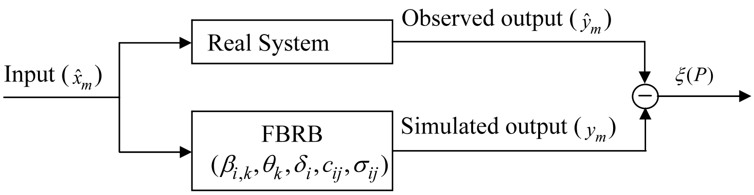

Fig. 1 Illustration of optimal learning process

attribute weights δ_i may have significant impact on the performance of a belief rule-based system. Moreover, the forms of fuzzy membership functions in the antecedent of a rule are an important factor for the system performance. θ_k, δ_i, the parameters c_{ij} and σ_{ij} of fuzzy membership functions therefore should be trained for achieving desirable performances if data is available.

In this section, optimization models and procedures are investigated in the *FUEBER* framework based on the optimization models for training general belief rules based systems (Yang et al. 2006c), in order to help generate desirable fuzzy belief rules and other system parameters simultaneously. Figure 1 shows the process of training a belief rule base (Yang et al. 2006c) with additional elements to deal with fuzziness, where $\hat{x}_m$ is a given input, $\hat{y}_m$ the corresponding observed output, either measured using instruments or assessed by experts, y_m the simulated output generated by the belief rule based system, and $\xi(P)$ the difference between $\hat{y}_m$ and y_m, as defined later.

It is desirable that $\xi(P)$ is as small as possible where P is the vector of training parameters including $\beta_{i,k}, \theta_k, \delta_i, c_{ij}, \sigma_{ij}$. This objective is difficult to achieve if a FBRB is constructed using expert judgments only. An optimal learning method is used to adjust the parameters in order to minimize the difference between the observed output ($\hat{y}_m$) and the simulated output (y_m), or $\xi(P)$. Such an optimally trained FBRB may then be used to predict the behavior of the system. In general, the general optimal learning problem can be represented as the following nonlinear (multiple objective) programming problem (Yang et al. 2006c).

$$\begin{array}{ll} \min & f(\boldsymbol{P}) \\ \text{s.t.} & A(\boldsymbol{P}) = 0, \qquad B(\boldsymbol{P}) \geq 0, \end{array} \tag{9}$$

where $f(\boldsymbol{P})$ is the objective functions, $\boldsymbol{P}$ the training parameter vector, $A(\boldsymbol{P})$ the equality constraint functions and $B(\boldsymbol{P})$ the inequality constraint functions.

In the learning process, a set of observations on the system input and output is required. In the following, we assume that a set of observation pairs $(\boldsymbol{x}, \boldsymbol{y})$ is available, where $\boldsymbol{x}$ is an input vector and $\boldsymbol{y}$ the corresponding output vector. Both $\boldsymbol{x}$ and $\boldsymbol{y}$ can be either numerical, judgmental or both. The format of the objective functions is important for the parameter optimization. Depending on the type of input and output, the optimal learning model can be constructed in different ways, as discussed in details in the following subsections.

In summary, due to the use of fuzzy belief rules, the *RIMER* (or *FURBER* in a fuzzy environment) method provides an analytical description of relationships between system's inputs and outputs. This enables a belief rule-based system to act as a generic functional mapping from system inputs to outputs and allows powerful learning techniques to be used for parameter training and system updating on the basis of both numerical data and human judgments, this is one of the main feature in the training models.

3.2 The optimization model based on the numerical output form

In this case, a training set (TS) composed of M input-output pairs (x_m, y_m) $(m = 1, \dots, M)$, where y_m is a numerical data (e.g., *risk rank index value*).

For instance, in real applications, the value of overall risk ranking index associated with various causes to a technical failure may be required, on the basis of which the identified potential causes can be ranked. The ranking results for risk due to various potential causes may help designers understand the anticipated technical problem in question so that an improved risk reduction measure can be incorporated in the new design or a more innovative design can be carried out to reduce the potential risk as estimated (Sii and Wang 2002).

The output in (7) gives an overall distribution. The risk ranking index associated with the cause to a failure can also be generated from the distribution if required. To calculate risk ranking index values associated with various causes to a technical failure, it is required to describe the safety expressions $\{D_1, D_2, \dots, D_N\}$, for example, {*Good, Average, Fair, Poor*} using numerical values, i.e., the utility of each linguistic safety expression. The utility values associated with the defined safety expressions can be assigned by experts or can be trained as shown in this paper. Suppose the utility (or score) of a safety expression D_j is denoted by $u(D_j)$ $(j = 1, \dots, N)$, and C is the anticipated cause to a technical failure. The expected utility (or score) of C is given as follows (Yang 2001):

$$u(C) = \sum_{j=1}^{N} u(D_j)\beta_j. \tag{10}$$

Note these utilities are used for characterizing an assessment but not used in the aggregation process. $u(D_j)$ can be estimated using the decision maker's preferences. Without loss of generality, suppose that the least preferred grade having the lowest utility is D_1 and the most preferred grade having the highest utility is D_N, i.e., $0 \le u(D_i) < u(D_j) \le 1$ if $i < j$; $i, j = 1, \dots, N$. Therefore, the utility of the individual consequent (D_j) should be determined. That means the additional parameter, utilities, should also be tuned alongside the belief matrix and the weights.

Hence, in the case of numerical output given in the training data set, the output of the system is given by (7), i.e., a numerical score. Therefore, the objective of the tuning method is to determine the belief matrix $((\boldsymbol{\beta})_{N\times L})$, the rule weights, the attribute weights $(\boldsymbol{\delta}, \boldsymbol{\theta})$, the central value $\boldsymbol{C}$, the variance $\boldsymbol{\sigma}$, and the utilities $u(D_j)_j$, which are denoted by $\boldsymbol{P}$. The objective function is to minimize the mean square error (MSE) criterion defined as follows:

$$\min_{\boldsymbol{P}}\{\xi(\boldsymbol{P})\}, \tag{11}$$

where, $\xi(P) = \frac{1}{M}\sum_{m=1}^{M}(y_m - \hat{y}_m)^2$, $y_m = \sum_{j=1}^{N} u(D_j)\beta_j(m)$ is the expected utility (or score) of the output corresponding to the given input $\boldsymbol{a}_m^*$ for the mth input vector in a training set $(m = 1, \dots, M)$, $\beta_j(m)$ is given by (8) for the mth input in the training set $(m = 1, \dots, M)$. M is the number of points in the training set, $\hat{y}_m$ is the measured confidence score or the expected output. $(y_m - \hat{y}_m)$ is the residual at the mth point. The constraint conditions for a scrip belief rule base are given as follows (Yang et al. 2006c):

(a) A belief degree (subjective probability) must not be less than zero or more than one

$$0 \le \beta_{jk} \le 1, \quad j = 1, \dots, N;\ k = 1, \dots, L. \tag{11a}$$

(b) If the kth belief rule is complete, its total belief degree in the consequent will be equal to one

$$\sum_{j=1}^{N} \beta_{jk} = 1. \tag{11b}$$

Otherwise, the total belief degree is less than one.

(c) A rule weight is normalized so that it is between zero and one

$$0 \le \delta_i \le 1, \quad i = 1, \ldots, 3. \tag{11c}$$

(d) An attribute weight is normalized so that it is between zero and one

$$0 \le \theta_k \le 1, \quad k = 1, \ldots, L. \tag{11d}$$

(h) For qualitative output, the score (utility) of a consequent can be normalized so that it is between zero and one

$$0 \le u(D_j) \le 1; \quad j = 1, \ldots, N. \tag{11e}$$

(i) The more preferred a consequent, the higher its score

$$u(D_i) < u(D_j) \quad \text{if } i < j;\, i, j = 1, \ldots, N. \tag{11f}$$

The following additional constraint conditions for a FBRB must also be satisfied.

(e) The boundary condition of the central value of the fuzzy membership function (Gaussian function)

$$lb_{MF} \le c_{ij} \le ub_{MF}, \quad i = 1, \ldots, 3;\; j = 1, \ldots, J_i. \tag{11g}$$

lb_{MF} and ub_{MF} are the bounds of the universal course of fuzzy membership functions.

(f) The monotonicity condition of the central value of the fuzzy membership function (Gaussian function)

$$c_{ij} \le c_{ik} \quad \text{if } j \le k,\, j, k = 1, \ldots, J_i. \tag{11h}$$

(g) The boundary condition of the variance at the central value of the fuzzy membership function

$$0 \le \sigma_{ij}, \quad i = 1, \ldots, 3;\; j = 1, \ldots, J_i. \tag{11i}$$

Therefore, all generations of the optimization algorithm are used to get the minimal mean square error. Equation (11) is a multi-variable constrained nonlinear single-objective optimization problem and can be solved using existing nonlinear optimization software packages. In this paper, the Matlab Optimization Toolbox is used (Coleman et al. 1999). Since the function to be minimized and the constraints are all continuous, this optimization problem can be solved using the optimization function FMINCON provided in MATLAB. FMINCON is a function to find a minimum of a constrained nonlinear multivariable function starting at an initial estimate. The basic format for using FMINCON is

$$\boldsymbol{P} = \text{fmincon}(\text{fun}, \boldsymbol{P}_0, \mathbf{A}, \mathbf{b}, \mathbf{Aeq}, \mathbf{beq}, \mathbf{lb}, \mathbf{ub}, \mathbf{nonlcon}), \tag{12}$$

where $y = \mathbf{fun}(\mathbf{x})$ is the function to be minimized, $\boldsymbol{P}_0$ is the initial guess, $\mathbf{A} * \mathbf{P} \le \mathbf{b}$ are linear inequality constraints, $\mathbf{Aeq} * \mathbf{P} = \mathbf{beq}$ are linear equality constraints, $\mathbf{lb}$ and $\mathbf{ub}$ are lower and upper bounds for each variable (if available), and $[\mathbf{c}, \mathbf{ceq}] = \mathbf{nonlcon}\ (\mathbf{P})$ is a function containing the nonlinear inequalities $\mathbf{c} \leq 0$ and nonlinear equalities $\mathbf{ceq} = 0$. Use [ ] for parameters which are not present in the problem. In the current problem, the objective function $\mathbf{fun} = \xi$ is given in (11). The tuning parameters are defined into a matrix $\boldsymbol{P}$. $\boldsymbol{P}_0$ is the initial guess of the parameter matrix. There is no nonlinear constraint in this model. The problem is re-arranged into the standard form required in MATLAB, and the FMINCON function is then used to solve the problem.

3.3 The optimization algorithm based on the output in the form of the subjective judgment

In this case, a training set composed of M input-output pairs (x_m, y_m) $(m = 1, \dots, M)$, where y_m can be a subjective judgment, i.e., a distributed assessment on the linguistic value with belief. Notice that the single judgment as one linguistic value can be regarded as a special case of the distribution assessment.

The output of *FURBER* is actually a distribution safety assessment instead of a single numerical score, which provides a panoramic view about the output status, from which one can see the variation between the original output and the revised output on each linguistic term. A distribution is easy to understand and flexible to represent output information than a single average value. Especially it is very useful in the case that the output is difficult to quantify. For example, the linguistic terms in safety (e.g., *good*, *fair*, or *poor*, etc.) have no clearly defined bases and are difficult to define quantitatively. So, it is better to draw a conclusion using the same linguistic terms as the ones in the consequent but with different degrees of belief. Such subjective judgments with belief are useful in such areas as safety/risk classification.

The previous optimization problem has a single objective function. This model is based on the output in the form of the subjective judgment for solving problems with multi-objective functions using Fminimax in Matlab (Coleman et al. 1999) while the output is given as a belief distribution. The final conclusion generated by aggregating the L rules shown in (7), which are activated by the actual input, is represented as follows:

$$\{(D_j, \beta_j), j = 1, \dots, N\}. \tag{13}$$

So each (D_j, β_j) $(j \in \{1, \dots, N\})$ can be regarded as one component of the multi-objective function vector. The function to be minimized is continuous. This multi-objective optimization problem is solved using the Fminimax function in MATLAB referred to as the *minimax problem* defined as follows (Yang et al. 2006c):

$$\begin{aligned} &\min_{\boldsymbol{Q}} \max_{\{\xi_j\}} \{\xi_j(\boldsymbol{Q}); j = 1, \dots, N\} \\ &\text{s.t.} \quad \text{(11a)–(11d) and (11g)–(11i)}, \end{aligned} \tag{14}$$

where

$$\xi_j(\boldsymbol{Q}) = \frac{1}{M} \sum_{m=1}^{M} (\beta_j(m) - \hat{\beta}_j(m))^2, \quad j = 1, \dots, N. \tag{14a}$$

$\boldsymbol{Q}$ is the tuning parameter vector. $\beta_j(m)$ is given by (8) for the mth input in a training set. M is the number of points in the training set, $\hat{\beta}_j(m)$ is the observed belief corresponding to the individual consequent D_j. $(\beta_j(m) - \hat{\beta}_j(m))$ is the residual at the mth point. The tuning parameters are beliefs, weights and parameters of fuzzy membership functions, without utilities. The constraints are (11a)–(11d) and (11g)–(11i). Equation (14) is an N-objective and multi-variable nonlinear optimization problem.

The minimax method (Coleman et al. 1999; Miettinen 1999; Marler and Arora 2004) is to minimize a worst case objective function. In other words, the purpose of the minimax formulation strategy is to minimize the maximum relative deviation of the objective function from its minimum objective function value. The objective functions are assigned weights, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$ indicating the designer's subjective preferences, with $0 \le \omega_j \le 1$ and $\sum_{j=1}^{N} \omega_j = 1$. Suppose ξ_j^+ and ξ_j^* are the feasible maximum and minimum values for the jth objective .Then, the objective function can be formulated as follows:

$$\min_{\boldsymbol{Q}} \max_{j=1,\ldots,N} \{\varphi_j(\boldsymbol{Q})\} \tag{15}$$

with

$$\begin{aligned} &\varphi_j(\boldsymbol{Q}) = \frac{\xi_j(\boldsymbol{Q}) - \xi_j^*}{\xi_j^+ - \xi_j^*}, \\ &\xi_j^+ - \xi_j^* > 0, \quad j = 1, \ldots, N. \end{aligned} \tag{15a}$$

The computational steps of the minimax method are summarized below:

Step 1. Based on the multi-objective function, solve the following single-objective optimization problem with the conventional methods individually (e.g., FMINCON).

$$\begin{aligned} &\min_{\boldsymbol{Q}} \xi_j(\boldsymbol{Q}) \\ &\text{s.t.} \quad \text{(11a)–(11d) and (11g)–(11i)}, \end{aligned} \tag{16}$$

where $j = 1, \ldots, N$, $\xi_j(\boldsymbol{Q})$ is given by (15a), and there are N single-objective optimization problems to solve. Suppose ξ_j^* is the optimal solution $\{j = 1, \ldots, N\}$.

Step 2. Suppose a relative weight vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$ **is provided**. In this paper all objectives are assumed to have equal weights.

Step 3. The multi-objective optimization problem can be formulated as follows:

$$\begin{aligned} &\min_{\boldsymbol{Q}} \max_{j=1,\ldots,N} \left\{ \omega_j \cdot \frac{\xi_j(\boldsymbol{Q}) - \xi_j^*}{\xi_j^+ - \xi_j^*} \right\} \\ &\text{s.t.} \quad \text{(11a)–(11d) and (11g)–(11i)}. \end{aligned} \tag{17}$$

Step 4. Arrange the problem into the standard form in MATLAB, and use the MATLAB function, FMINIMAX, to solve it.

The optimization processes proposed in subsections Sects. 3.2–3.3 are a kind of iterative process, which continues until the mean square error becomes smaller than the specified tolerance.

3.4 The optimization model based on the input-output in form of interval

In real applications, data sets used to train a belief rule base may not be precise but within a certain range. In this case, entries of input are within a certain range, or an interval vectors. Moreover, target output may not contain only points, but also intervals. So the training set is composed of M input-output pairs $(\hat{x}_m, \hat{y}_m)$ $(m = 1, \ldots, M)$, where $\hat{x}_m$ and $\hat{y}_m$ are interval vectors. So it is the optimization algorithm based on the input-output in form of intervals.

For the optimization, a three-step optimization model may be used. Suppose that $\xi(\boldsymbol{P})$ is the objective function defined in (9) and $\boldsymbol{P}$ is the parameters to be trained.

The optimization requires the output set $\xi(\hat{x}_m, \boldsymbol{P}) \supseteq \hat{y}_m$ for $m = 1, \dots, M$ with a minimum overlap. So the problem is both to minimize the volume of the output $\xi(\hat{\boldsymbol{x}}_m, \boldsymbol{P})$ and to minimize the distance between the center of $\xi(\hat{\boldsymbol{x}}_m, \boldsymbol{P})$ and $\hat{y}_m$. Therefore, the problem is formulated as follows.

$$\min_{\boldsymbol{P}} \sum_{m=1}^{M} (\bar{y}_m - \underline{y}_m)^2, \tag{18}$$

$$\min_{\boldsymbol{P}} \sum_{m=1}^{M} \left[\left(\frac{\bar{y}_m - \underline{y}_m}{2} \right) - \operatorname{mid} \hat{y}_m \right]^2, \tag{19}$$

where $\operatorname{mid} \hat{y}_m$ is the center of the interval $\hat{y}_m$. $\bar{y}_m$ is the optimal point such that

$$\max_{\boldsymbol{x}_m \in \boldsymbol{X}_m} y_m(x_m, \boldsymbol{P}_0), \tag{20}$$

$\underline{y}_m$ is the optimal point such that

$$\min_{\boldsymbol{x}_m \in \boldsymbol{X}_m} y_m(x_m, \boldsymbol{P}_0), \tag{21}$$

where $y_m = \sum_{j=1}^{N} u(D_j)\beta_j(m)$ is defined in (9), and $\boldsymbol{P}_0$ is the initial training parameter vector.

By combining (18) and (19) with an index λ, the objective function can be written as

$$\min_{\boldsymbol{P}} \lambda \sum_{m=1}^{M} (\bar{y}_m - \underline{y}_m)^2 + (1-\lambda) \min_{\boldsymbol{P}} \sum_{m=1}^{M} \left[\left(\frac{\bar{y}_m - \underline{y}_m}{2} \right) - \operatorname{mid} \hat{y}_m \right]^2. \tag{22}$$

The index λ may be determined by the user according to the importance of the objective functions (18) and (19) in an application.

Notice that besides the constraints mentioned in (11a)–(11i), the additional constraints are given by

$$\xi(\hat{x}_m, \boldsymbol{P}) \supseteq \hat{y}_m \quad \text{for } m = 1, \dots, M. \tag{23}$$

In the above optimal models proposed in subsections Sects. 3.2–3.4, a fuzzy belief rule-based system can be initialized either by experts or arbitrarily, depending on initial information available. In the former case, the domain knowledge of experts is used to assign the initial belief-rule-base matrix, the weights, and the parameters of the fuzzy membership function, which are then refined using the above models if more information becomes available. In the latter case, the initial belief rule-base matrix, the weights, and the parameters of the fuzzy membership function are arbitrarily generated and then refined using the above models. However, a trained belief rule base with arbitrary initialization may generate intuitively wrong conclusions if not all rules are trained, which may be the case if the observed data set does not cover all possible regions the belief rule base is designed to operate (Yang et al. 2006c).

4 A numerical study

The *FURBER* framework (Liu et al. 2004) has been applied to modelling system safety of an offshore and marine engineering system: a floating production storage offloading (FPSO) system (Chen and Moan 2002; McCaul 2001), with special focuses on collision risk between FPSO and a shuttle tanker due to technical failures during a tandem offloading operation.

4.1 Problem description and the optimization models

This study concentrates on the risk evaluation of the major hazards threatening FPSO overall rather than focusing on specific areas of design. The investigated FPSO system is a turret-moored system connected through flexible risers to remote sub sea wells. The safety assessment provides a means for screening the safety implications which would influence the development of the concept. This will permit these safety and economic consequences of the concept to be considered in the early design processes and also highlights areas where little guidance or partial experience was available, especially with these types of innovative developments.

In this section, safety assessment is carried out on risk introduced by the collision of FPSO and a shuttle tanker during a tandem offloading operation. Only technical failures caused risk is assessed here, though the operational failure has been also recognised as one of the major causes of collisions.

According to the literature survey, the technical failures that might cause collisions between FPSO and a shuttle tanker during tandem offloading operations are malfunction of propulsion systems (Chen and Moan 2002). The four major causes to the technical failures are:

(1) Controllable pitch propeller (CPP) failure;
(2) Thruster failure;
(3) Position reference system (PRS) failure;
(4) Dynamics positioning system failure (DP).

An example within the same application framework as analyzed in Liu et al. (2004) is used to demonstrate how the optimization method can be implemented in safety analysis. For illustration purpose, we only consider the safety assessment related to a controllable pitch propeller (CPP) failure to demonstrate the procedure involved in the optimization of belief rule base and other knowledge representation parameters in *FURBER*.

Twenty-seven fuzzy rules from a total of 245 rules (Sii and Wang 2002; Liu et al. 2004) are extracted and used in our example, which are described in Appendix and used as the initial belief rule-base in the learning process. The linguistic terms (*average*, *frequent*, *highly frequent*) are defined for **FR**, (*moderate*, *critical*, *catastrophic*) for **CS**, and (*likely*, *highly likely*, *definite*) for FCP respectively. Because 3 antecedent attributes FR, CS, and FCP are each defined by 3 linguistic terms, the number of rules in the rule-base is $3\times3\times3=27$. The detailed definitions of their linguistic terms can be found in Sii and Wang (2002), Liu et al. (2004), and the corresponding initial membership functions are given by the Gaussian function as shown in Fig. 3 in Appendix.

It is assumed that each input parameter may be fed into the proposed safety model in a single deterministic value although there are other input forms that can be used to address inherent uncertainty associated with the data as discussed in Sii and Wang (2002), Liu et al. (2004).

Note that the training data are given in two kinds of output forms. So the corresponding optimization formulation are given in (11) for a numerical output form and (14) for a

Table 2 Training data (input values, expected values of output, trained output value, system output value before training)

Training No.	FR	CS	FCP	Risk rank index		
				Expected	Trained	Before trained
1	4.75	8.25	7.6	0.398	0.3998	0.2057
2	9	8.5	9	0.263	0.2626	0.2002
3	5	4.5	7	0.458	0.4576	0.5990
4	7	5	9.4	0.482	0.4819	0.5075
5	6.5	8	7.5	0.405	0.4031	0.2086
6	7.15	7.95	7.25	0.368	0.3688	0.2192
7	7	9	7	0.439	0.4386	0.2300

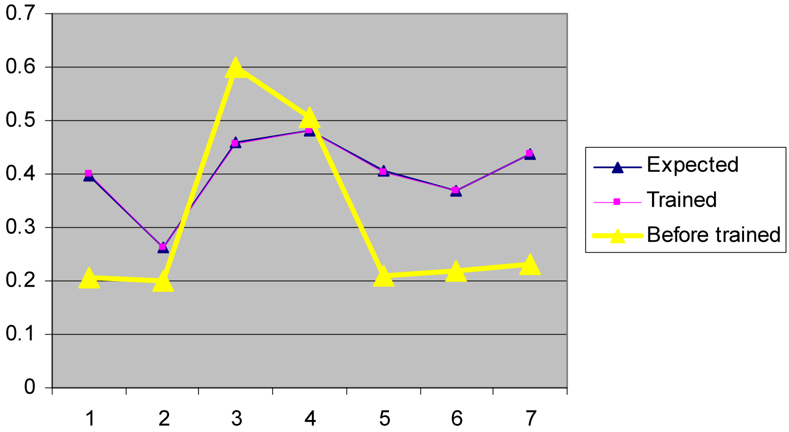

Fig. 2 The comparison of result on the training set

subjective judgment form respectively, where $L = 27$, $T = 3$ and $N = 4$. The computation steps are given in Sect. 3.

The existing optimization tool (i.e. FMINCON function in MATLAB) is used to solve problem (11). In this example, the calculation is based on a medium-scale algorithm in FMINCON, and all the results are based on the Sequential Quadratic Programming (SQP), Quasi-Newton, and line-search method in FMINCON. The FMINIMAX function in MATLAB is used to solve problem (14) in this paper.

4.2 Experiment results base on the numerical output

To train the belief rule base for the above example, 14 sets of data are used. We would like to tune and refine beliefs, weights, parameters of fuzzy membership functions, and the utilities of the linguistic terms of the safety estimate. The available data are partitioned into a training set and a test set. The training set is used for parameter training purposes. Once a desired performance measure value is achieved, the corresponding approximation error on the test data is measured.

In this example, we use 7 sets of training data for the parameter estimation (shown in Table 2), where the original input and output data were provided as numerical numbers. The initial belief matrix is supposed to be given by Experts as shown in Appendix (Liu et al.

Table 3 Test data (input values, expected values of output, tested output value)

Test No.	FR	CS	FCP	Risk rank index	
				Expected	Tested
1	7.5	8	6	0.448	0.5501
2	7.5	7.2	7.1	0.353	0.3226
3	7.5	6	7	0.388	0.4835
4	9.2	7	6	0.43	0.5743
5	8	8.5	5	0.401	0.4637
6	5	6.75	7	0.489	0.4407
7	7.95	8.25	7.9	0.323	0.2708

2004). The initial attribute weights and the rule weights are all set to 1. The training results are compared in Table 2 and Fig. 2 as well.

As shown in Table 2, there is a big difference between the initial and expected outputs. The big difference is due to the fact that the initial outputs are obtained before the belief rule matrix and the weights are tuned. The difference is not huge because of the small size of the rule-base. However, it would be difficult to obtain such a relatively accurate belief rule matrix without training for a complex system. As shown in Fig. 2, after the parameter optimization, the performance of the system is greatly improved and the trained belief rule based system can simulate the real system with great accuracy, which is strongly nonlinear. This investigation shows that the belief-rule-base matrix and other knowledge representation parameters can be subjectively assigned initially using partial human knowledge and then refined by training if input-output data become available. The optimized belief rule base is given in Table 6 of Appendix.

The initial fuzzy membership function and the optimized fuzzy membership function are shown in Figs. 3 and 4 of Appendix respectively. The initial utility of the linguistic term of the safety estimate is given as $u(\text{good}) = 1$, $u(\text{average}) = 0.8$, $u(\text{fair}) = 0.6$, $u(\text{poor}) = 0.2$. The optimized utility is obtained as: $u(\text{good}) = 0.9987$, $u(\text{average}) = 0.8559$, $u(\text{fair}) = 0.1769$, $u(\text{poor}) = 0.0501$.

After the tuning, a testing data consisting of 7 records (Table 3) is used. The expected and actual values of the output are listed in Table 3.

One may notice that the testing results in the above table indicate that 90% of them are already correct within the specified tolerance 0.1 (except test 4). The error including test 4 would be decreasing with more training data available.

4.3 Experiment results based on judgment assessments

To validate the trained model in this case, the available data are partitioned into a training data set and a test data set. The first 9 sets of data, as shown in Table 4, are used as the training data for parameter estimation. After training, the other 5 sets of data, as shown in Table 5, are used for validation.

Table 4 shows the comparison between the trained results and the expected distributed assessments generated on the basis of the beliefs from experts. In this experimental example, the error tolerance for each single objective minimization is set to 1.0e-5 and the maximum iteration is set to 100, and the error tolerance for a multi-objective minimization is set to 1.0e-4 and the maximum iteration is set to 100 to avoid dead loops in the learning process. The initial belief rule-base is given by experts shown in Appendix. The initial fuzzy membership function and the optimized fuzzy membership function are shown in Figs. 3 and 5 in Appendix respectively. The optimized belief rule base is given in Table 7 of Appendix.

Table 4 Trained results comparison based on judgment observations

Train No.	Antecedents			Safety estimate (Distribution assessment with belief)							
	FR	CS	FCP	Expected				Trained			
				Good	Average	Fair	Poor	Good	Average	Fair	Poor
1	7.75	8.25	7.6	0	0.0123	0.3641	0.6236	0.0289	0.0289	0.2908	0.6514
2	7	8	7.25	0	0.0033	0.3090	0.6876	0.0526	0.0526	0.3054	0.5893
3	8	8.5	7	0	0.0057	0.3735	0.6208	0.0251	0.0252	0.2893	0.6604
4	7	7	5.5	0	0.0373	0.7802	0.1825	0.0062	0.0062	0.7566	0.2309
5	6.5	8	7.5	0	0.0640	0.4165	0.5195	0.0617	0.0617	0.3200	0.5567
6	7.15	7.95	7.25	0	0.0013	0.4179	0.5808	0.0462	0.0462	0.3865	0.5210
7	7	8.5	7	0	0.0047	0.6151	0.3802	0.0692	0.0692	0.5492	0.3125
8	7.5	8	6	0	0.0041	0.6142	0.3817	0.0158	0.0158	0.5579	0.4105
9	7.5	7.2	7.1	0	0.0080	0.3694	0.6226	0.0282	0.0282	0.3084	0.6351

Table 5 Test results comparison base on judgment observations

Test No.	Antecedents			Safety estimate (Distribution assessment with belief)							
	FR	CS	FCP	Expected				Tested			
				Good	Average	Fair	Poor	Good	Average	Fair	Poor
1	7.5	8.5	7	0	0.0102	0.3595	0.6303	0.0409	0.0409	0.3334	0.5848
2	7	7	6	0	0.0097	0.6926	0.2977	0.0155	0.0155	0.6172	0.3518
3	8	8.5	7.5	0	0.0097	0.3930	0.5973	0.0244	0.0245	0.3053	0.6458
4	7.25	6.75	7	0	0.0200	0.5733	0.4067	0.0254	0.0254	0.4712	0.4781
5	7.95	8.25	7.9	0	0.0256	0.2688	0.7056	0.0233	0.0233	0.3080	0.6455

The trained belief rule-based system is then used to generate outputs for the test input data. The trained outputs in terms of belief degrees to the linguistic terms are very close to the observed outputs. This shows the capability of the belief rule based system to simulate real system using judgmental information (output in this case). The test result is shown in Table 5. The mean square error is 0.01. So the test performance is good as well.

5 Discussions

As analyzed in detail in the earlier study (Yang et al. 2006c), some further discussions about this study are summarized as follows.

(1) The above study shows that after an initial belief rule based system is established using expert knowledge its performance can be further improved if partial input-output data is available. With the accumulation of new knowledge, the system can be progressively trained and updated to better mimic a real system. Note that the training data in the above example do not cover the full range of regions where the belief rule base is designed to operate. For example, Rules 19 and 21 in Table 6 were not trained at all as the belief degrees of these rules remain unchanged after the training, in whichever ways

the belief rule base was initialized, as marked in bold in Table 6. When activated, such rules untouched during training could lead to irrational conclusion if they were initially assigned randomly or without care. For example, if the belief rule base is randomly built initially, then the three rules would lead to intuitively wrong conclusions if they were later activated after the training of the rule base.

(ii) The objective functions in the optimization models investigated in this paper are continuous and differentiable nonlinear and the constraint functions are all linear. So these optimization models can be readily solved using existing nonlinear optimization software such as Matlab optimization tool box, Lindo, and Excel solver. Also note that the optimization models may have multiple optima and using expert knowledge for initialization could help to achieve a trained belief rule base closer to expert expectations, leading to better results in extrapolation. We believe that this paper reports the progress of a research programme towards providing a global optimization scheme in a long run.

(iii) The final performance of a supervised belief rule-based learning system depends on both its system structure and system parameters. However, to generate globally optimal system performance in general circumstances would also require the development of global optimization models. If a global optimization model is proposed in future with the structure of a belief rule base subject to change, then it would be beneficial to design hybrid algorithms by combining mathematical programming techniques with adaptive search or heuristic methods such as Genetic Algorithm and Simulated Annealing.

6 Conclusion

Starting from a fuzzy belief rule base (FBRB) for engineering system safety analysis, the optimization training models were investigated in this paper and used to fine tune the parameters of a fuzzy belief rule based system initially built using human knowledge. This optimization training method provides a practical support to construct flexible and reliable belief rule bases, which can optimally imitate complex reasoning processes and represent nonlinear or non-smooth relationships using both human knowledge and numerical data.

Inheriting the unique features of the optimal training models developed for general belief rule base (Yang et al. 2006c), the major feature of the optimal models for FBRB is that only partial input and output information is required, which can be either incomplete or vague, either numerical (deterministic value or an interval) or judgmental. The models can be used to fine tune a *FBRB* whose internal structure is initially decided by experts' domain specific knowledge or common sense judgments. Conclusions drawn from such a trained *FBRB* with partially built-in expert knowledge can simulate real situations in a meaningful and locally optimal way. This may allow the fuzzy belief rule-based system to be equipped with the self-tuning capability of simulating an engineering system safety analysis process.

The validation of training results needs to be extensively investigated for larger systems. The method investigated in this paper can provide a non-black box simulator to support engineering system safety analysis at large.

Acknowledgements This work was partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant No.: GR/S85498/01 and Grant No.: GR/S85504/01, as well as the Research Project TIN2006-02121. The authors are grateful to the two anonymous referees for their constructive comments that have helped to improve the quality of the paper to its current standard.

Appendix

The RULE BASE use in the current case study of this paper is originally extracted from the rule-base established in Sii and Wang (2002)

1. Rule # 137: *IF the* ***failure rate*** *is* *average* *AND the* ***consequence severity*** *is* *catastrophic* *AND the* ***failure consequence probability*** *is* *likely* *THEN the* ***safety estimate*** *is* *fair*
2. Rule # 139: *IF the* ***failure rate*** *is* *average* *AND the* ***consequence severity*** *is* *catastrophic* *AND the* ***failure consequence probability*** *is* *highly likely* *THEN the* ***safety estimate*** *is* *poor*
3. Rule # 140: *IF the* ***failure rate*** *is* *average* *AND the* ***consequence severity*** *is* *catastrophic* *AND the* ***failure consequence probability*** *is* *definite* *THEN the* ***safety estimate*** *is* *poor*
4. Rule # 193: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *moderate* *AND the* ***failure consequence probability*** *is* *likely* *THEN the* ***safety estimate*** *is* *fair*
5. Rule # 195: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *moderate* *AND the* ***failure consequence probability*** *is* *highly likely* *THEN the* ***safety estimate*** *is* *fair*
6. Rule # 196: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *moderate* *AND the* ***failure consequence probability*** *is* *definite* *THEN the* ***safety estimate*** *is* *poor*
7. Rule # 200: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *critical* *AND the* ***failure consequence probability*** *is* *likely* *THEN the* ***safety estimate*** *is* *fair*
8. Rule # 202: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *critical* *AND the* ***failure consequence probability*** *is* *highly likely* *THEN the* ***safety estimate*** *is* *poor*
9. Rule # 203: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *critical* *AND the* ***failure consequence probability*** *is* *definite* *THEN the* ***safety estimate*** *is* *poor*
10. Rule # 207: *IF the* ***failure rate*** *is* *frequent* *AND the* ***consequence severity*** *is* *catastrophic* *AND the* ***failure consequence probability*** *is* *likely* *THEN the* ***safety estimate*** *is* *poor*

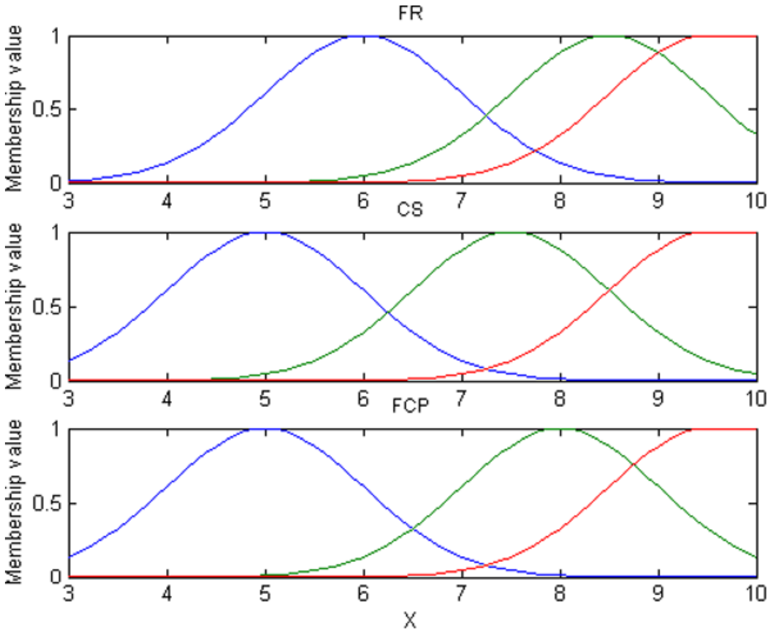

Fig. 3 The initial fuzzy membership function

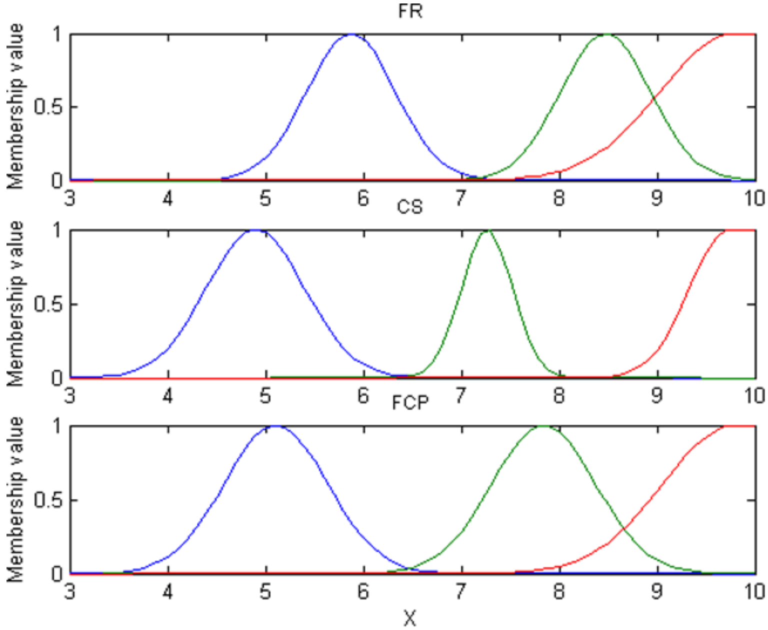

Fig. 4 The optimized fuzzy membership function for the case based on the numerical output

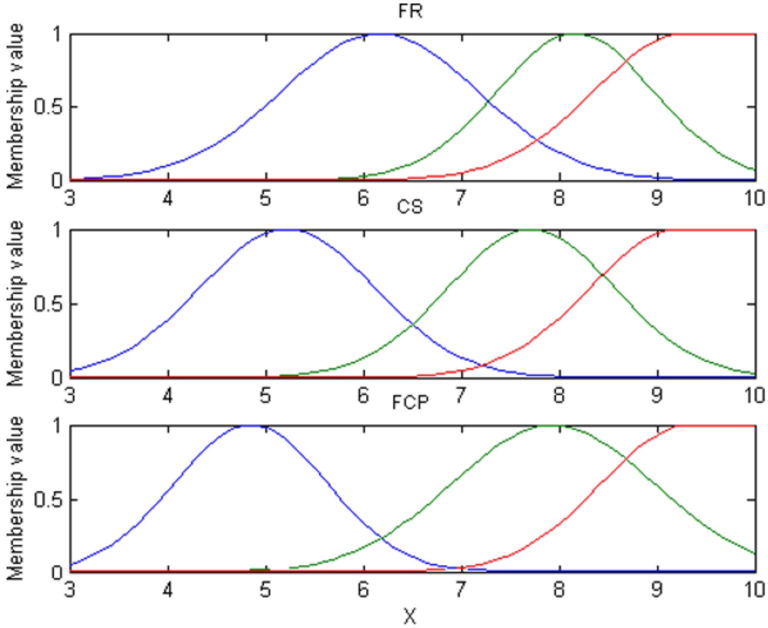

Fig. 5 The optimized fuzzy membership function for the case based on the subjective output

Table 6 The optimized belief rule base for the case based on the numerical output

Rule No.	Trained belief structure of the rule			
	Good	Average	Fair	Poor
1	0.0020	0.0001	0.9434	0.0545
2	0.0155	0.0250	0.5058	0.4537
3	0.0013	0.0003	0.8364	0.1620
4	0.0108	0.0038	0.9852	0.0001
5	0.2144	0.1123	0.0230	0.6503
6	0.0004	0.0000	0.0000	0.9996
7	0.0604	0.0073	0.9323	0.0000
8	0.2699	0.1107	0.0256	0.5939
9	0.0005	0.0000	0.0000	0.9995
10	0.0000	0.0000	0.9996	0.0004
11	0.0002	0.0002	0.9820	0.0177
12	0.0002	0.0015	0.0002	0.9981
13	0.0231	0.0010	0.9717	0.0042
14	0.0323	0.0135	0.0043	0.9499
15	0.0635	0.0215	0.0163	0.8986
16	0.0537	0.0186	0.0002	0.9274
17	0.2584	0.1029	0.0013	0.6374
18	0.0916	0.0391	0.0033	0.8660
19	**0.0000**	**0.0000**	**1.0000**	**0.0000**
20	**0.0000**	**0.0001**	**0.9977**	**0.0022**
21	**0.0000**	**0.0000**	**0.0000**	**1.0000**
22	0.0030	0.0001	0.9959	0.0010
23	0.0450	0.0159	0.0010	0.9381
24	0.0796	0.0312	0.0011	0.8882
25	0.0066	0.0025	0.0000	0.9908
26	0.0780	0.0271	0.0009	0.8941
27	0.0865	0.0287	0.0003	0.8845

11. Rule # 209: *IF the* ***failure rate*** *is frequent AND the* ***consequence severity*** *is catastrophic AND the* ***failure consequence probability*** *is highly likely THEN the* ***safety estimate*** *is poor*
12. Rule # 210: *IF the* ***failure rate*** *is frequent AND the* ***consequence severity*** *is catastrophic AND the* ***failure consequence probability*** *is definite THEN the* ***safety estimate*** *is poor*
13. Rule # 214: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is negligible AND the* ***failure consequence probability*** *is likely THEN the* ***safety estimate*** *is fair*
14. Rule # 216: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is negligible AND the* ***failure consequence probability*** *is highly likely THEN the* ***safety estimate*** *is fair*
15. Rule # 217: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is negligible AND the* ***failure consequence probability*** *is definite THEN the* ***safety estimate*** *is fair*
16. Rule # 221: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is marginal AND the* ***failure consequence probability*** *is likely THEN the* ***safety estimate*** *is fair*
17. Rule # 223: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is marginal AND the* ***failure consequence probability*** *is highly likely THEN the* ***safety estimate*** *is fair*
18. Rule # 224: *IF the* ***failure rate*** *is highly frequent AND the* ***consequence severity*** *is marginal AND the* ***failure consequence probability*** *is definite THEN the* ***safety estimate*** *is fair*

Table 7 The optimized belief rule base for the case based on the judgment observations

Rule No.	Trained belief structure of the rule			
	Good	Average	Fair	Poor
1	0.0051	0.0051	0.9779	0.0120
2	0.0034	0.0034	0.0075	0.9856
3	0.0001	0.0001	0	0.9998
4	0.0074	0.0074	0.6440	0.3413
5	0.0589	0.0589	0.3851	0.4971
6	0.0134	0.0134	0	0.9731
7	0.0081	0.0081	0.9216	0.0622
8	0.1642	0.1642	0.3204	0.3511
9	0	0	0.0026	0.9974
10	0.0031	0.0031	0.0006	0.9932
11	0.0055	0.0055	0.0121	0.9769
12	0.0001	0.0001	0	0.9998
13	0	0	0.6501	0.3499
14	0.0347	0.0347	0.1171	0.8135
15	0.0168	0.0168	0.9004	0.0660
16	0.0126	0.0126	0.8564	0.1184
17	0.0277	0.0279	0.5970	0.3475
18	0.0094	0.0094	0.9474	0.0338
19	0	0	0.9972	0.0028
20	0	0	0.9938	0.0062
21	**0**	**0**	**0**	**1.0000**
22	0.0055	0.0055	0.8975	0.0915
23	0	0	0.0337	0.9663
24	0.0044	0.0044	0.0098	0.9814
25	0.0031	0.0031	0.0050	0.9888
26	0.0019	0.0019	0.0314	0.9648
27	0.0051	0.0051	0.9779	0.0120

19. Rule # 228: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*moderate*</u> *AND the* ***failure consequence probability*** *is* <u>*likely*</u> *THEN the* **safety estimate** *is* <u>*fair*</u>
20. Rule # 230: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*moderate*</u> *AND the* ***failure consequence probability*** *is* <u>*highly likely*</u> *THEN the* **safety estimate** *is* <u>*fair*</u>
21. Rule # 231: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*moderate*</u> *AND the* ***failure consequence probability*** *is* <u>*definite*</u> *THEN the* **safety estimate** *is* <u>*poor*</u>
22. Rule # 235: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*critical*</u> *AND the* ***failure consequence probability*** *is* <u>*likely*</u> *THEN the* **safety estimate** *is* <u>*fair*</u>
23. Rule # 237: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*critical*</u> *AND the* ***failure consequence probability*** *is* <u>*highly likely*</u> *THEN the* **safety estimate** *is* <u>*poor*</u>
24. Rule # 238: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*critical*</u> *AND the* ***failure consequence probability*** *is* <u>*definite*</u> *THEN the* **safety estimate** *is* <u>*poor*</u>
25. Rule # 242: *IF the* ***failure rate*** *is* <u>*highly frequent*</u> *AND the* **consequence severity** *is* <u>*catastrophic*</u> *AND the* ***failure consequence probability*** *is* <u>*likely*</u> *THEN the* **safety estimate** *is* <u>*poor*</u>
26. Rule # 244: IF the ***failure rate*** is <u>*highly frequent*</u> AND the ***consequence severity*** is

catastrophic AND *the* ***failure consequence probability*** is highly likely THEN the ***safety estimate*** is poor

27. Rule # 245: IF the ***failure rate*** is highly frequent AND the ***consequence severity*** is catastrophic AND the ***failure consequence probability*** is definite THEN the ***safety estimate*** is poor

In Table 6, RULES #1– #27 are trained rules corresponding to RULES #137– #245, for example, RULE #1 here is read as:

IF the ***failure rate*** *is* *average* *AND the* ***consequence severity*** *is* *catastrophic* *AND the* ***failure consequence probability*** *is* *likely* *THEN the* ***safety estimate*** *is* {(*Good*, 0.002), (*Average*, 0.0001), (*Fair*, 0.9434), (*Poor*, 0.0545)}

References

Bell, P. M., & Badiru, A. B. (1996). Fuzzy modelling and analytic hierarchy processing to quantify risk levels associated with occupational injuries—Part I: The development of fuzzy-linguistic risk levels. *IEEE Transactions on Fuzzy Systems*, *4*(2), 124–131.

Binaghi, E., & Madella, P. (1999). Fuzzy Dempster–Shafer reasoning for rule-based classifiers. *International Journal of Intelligent Systems*, *14*, 559–583.

Bowles, J. B., & Pelaez, C. E. (1995). Fuzzy logic prioritisation of failures in a system failure mode, effects and criticality analysis. *Reliability Engineering and System Safety*, *50*, 203–213.

Chen, H., & Moan, T. (2002). Collision risk analysis of FPSO-tanker offloading operation. In *21st International Conference on Offshore Mechanics & Arctic Engineering*, Oslo, Norway, 23–28 June 2002.

Coleman, T., Branch, M. A., & Grace, A. (1999). *Optimization toolbox—for use with Matlab*. The Mathworks Inc.

Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society, Series B*, *30*, 205–247.

Duckstein, L. (1994). Elements of fuzzy set analysis and fuzzy risk. In H. P. Nachtnebel (Ed.), *Decision support systems in water resources management* (pp. 410–430). Paris: UNESCO Press.

Liu, J., Yang, J. B., Wang, J., & Sii, H. S. (2003). Review of uncertainty reasoning approaches as guidance for maritime and offshore safety-based assessment. *Journal of UK Safety and Reliability Society*, *23*(1), 63–80.

Liu, J., Yang, J. B., Wang, J., Sii, H. S., & Wang, Y. M. (2004). Fuzzy rule-based evidential reasoning approach for safety analysis. *International Journal of General Systems*, *33*(2–3), 183–204.

McCaul, J. R. (2001). Special report—floating production systems. *Oil & Gas Journal*, June, 11.

Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Struct Multidisc Optim.*, *26*, 369–395.

Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Boston: Kluwer Academic.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton: Princeton University Press.

Sii, H. S., & Wang, J. (2002). *Safety assessment of FPSOs—The process of modelling system safety and case studies*. Report of the project—The application of approximate reasoning methodologies to offshore engineering design (EPSRC GR/R30624 and GR/R32413), Liverpool John Moores University, UK.

Smets, P. (1988). Belief function. In P. Smets, E. H. Mamdani, D. Dubois & H. Prade (Eds.), *Non-standard logics for automated reasoning* (pp. 253–277). London: Academic Press.

Wang, J., Yang, J. B., & Sen, P. (1995). Safety analysis and synthesis using fuzzy set modelling and evidential reasoning. *Reliability Engineering and System Safety*, *47*(3), 103–118.

Wang, Y. M., Yang, J. B., & Xu, D. L. (2006). Environmental impact assessment using the evidential reasoning approach. *European Journal of Operational Research*, *174*(3), 1885–1913.

Yang, J. B. (2001). Rule and utility based evidential reasoning approach for multi-attribute decision analysis under uncertainties. *European Journal of Operational Research*, *131*, 31–61.

Yang, J. B., & Sen, P. (1994). A general multi-level evaluation process for hybrid MADM with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, *24*(10), 1458–1473.

Yang, J. B., & Singh, M. G. (1994). An evidential reasoning approach for multiple attribute decision making with uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics*, *24*(1), 1–18.

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

Yang, J. B., & Xu, D. L. (2002b). Nonlinear information aggregation via evidential reasoning in multi-attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, *32*(3), 376–393.

Yang, J. B., Liu, J., Wang, J., & Sii, H. S. (2006a). A generic rule-base inference methodology using the evidential reasoning approach—RIMER. *IEEE Transactions on Systems, Man, and Cybernetics*, *36*(2), 266–285.

Yang, J. B., Wang, Y. M., Xu, D. L., & Chin, K. S. (2006b). The evidential reasoning approach for MCDA under both probabilistic and fuzzy uncertainties. *European Journal of Operational Research*, *171*(1), 309–343.

Yang, J. B., Liu, J., Xu, D. L., Wang, J., & Wang, H. W. (2006c). Optimization models for training belief rule based systems. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, *37*(4), 569–585.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zadeh, L. A. (1975a). The concept of a linguistic variable and its applications to approximate reasoning. Part I. *Information Sciences*, *8*, 199–249.

Zadeh, L. A. (1975b). The concept of a linguistic variable and its applications to approximate reasoning. Part II. *Information Sciences*, *8*, 301–357.

Zadeh, L. A. (1975c). The concept of a linguistic variable and its applications to approximate reasoning. Part III. *Information Sciences*, *9*, 43–80.