Scholar Publishing Group
International Journal of Big Data Intelligent Technology
https://doi.org/10.38007/IJBDIT.2025.060203
ISSN 2790-0932 Vol. 6, Issue 2: 21-35

# *Oil and Gas Fire Risk Based on Fuzzy Fault Tree and Bayesian Network*

**Yao Hu[1, a*] Liguang Qiao[2,b] and Feng Gu[3,c]**
[1]*College of Civil Aviation Safety Engineering, Civil Aviation Flight University of China, Guanghan 618307, Sichuan, China*
[2]*College of Economics and Management, Civil Aviation Flight University of China, Guanghan 618307, Sichuan, China*
[3]*Logistics service company, Civil Aviation Flight University of China, Guanghan 618307, Sichuan, China*
[a]*hy18328072523@163.com*, [b]*q15008267098@163.com*, [c]*gf19880615@163.com*
[*]*Corresponding author*

***Keywords:*** Safety Engineering; Oil and Gas Fire; Fault Tree Analysis; Bayesian Network; Risk Probability

***Abstract:*** Oil and gas fires are frequent disasters that occur during oil and gas storage and transportation. In order to prevent the occurrence of oil and gas fires and reduce human and material losses, this paper studied the risk prediction of oil and gas fires based on the fusion of fuzzy fault tree analysis (FFTA) and Bayesian network (BN) algorithm. Firstly, by conducting statistical analysis of accidents, 9 causal factors leading to oil and gas fires were identified. A fuzzy fault tree was established with oil and gas fire accidents as the top event, management issues, oil product issues, protection failures, and hazardous states as intermediate events, and 9 causal factors as basic events. The logical network connections between the fault trees were utilized to organize the causal relationships between each causal factor, and mapping relationships were used to connect the fault tree with Bayesian networks. Finally, combined with expert evaluation, fuzzy processing was performed to obtain the probability of occurrence for each root node (causal factor), and then the prior probabilities of intermediate nodes and leaf nodes were obtained through logical operations and Bayesian network full probability formulas. Through backward inference analysis, it was found that the posterior probability of the intermediate event “dangerous state” was relatively high. Combined with sensitivity analysis, it can be concluded that the fundamental cause of the intermediate event “dangerous state” is the basic event “equipment damage”, and the fundamental cause of the intermediate event “protection failure” is the basic event “protective system failure”. Therefore, it is recommended to take measures such as regular maintenance and inspection, installation of fault monitoring and early warning systems, backup equipment, backup power supply, development of emergency plans and training, and regular drills and evaluations. These measures help to improve the reliability and stability of protective systems and reduce the potential risk of oil and gas fires.

## 1. Introduction

The safety of oil and gas storage and transportation is an important issue in the oil and gas industry, which is directly related to the safety of people's lives and property and the stability of social and economic development. In the process of storage and transportation, once an accident occurs, it may cause serious casualties, environmental pollution, and economic losses, bringing immeasurable impacts to society.

At present, the analysis of oil and gas fire mainly focuses on the qualitative analysis of the cause of the first accident. The research mainly explains the occurrence of oil and gas fires by identifying and describing the causes of oil and gas fires. Based on various risk accidents of oil and gas stations and yards, Fan Yong's research [1] deeply analyzed the safety issues of oil and gas fires from multiple aspects such as management, operation, operation and norms, and revealed the significant impact of management and operation errors on fire risks. Zhang Xinlin's research [2] discussed in depth the causes of oil and gas equipment failures, combined oil and gas fire prevention measures with big data technology, and created mobile storage and transportation equipment with equipment as storage and transportation point using the Internet platform, so as to achieve optimal management and maintenance of oil and gas storage and transportation process, thus improving the efficiency of risk prevention and control. Hosseinnia Davatgar B[3] evaluated the importance of safety barriers in the prevention of major accidents through the analysis of safety equipment in oil and gas accidents, focused on the application of risk-based analysis on the Goliat platform, and extended it to related safety systems to assess dynamic risks throughout the plant. Dimaio F[4] uses multi-state Bayesian networks to simulate and evaluate the functional performance of safety gates in oil and gas plants. It uses the multi-state Bayesian network model to analyze and explain the hazards involved in the relevant plant processes, and proposes methods to simplify the risks. Although these qualitative analysis methods can effectively explain various causes of oil and gas fires, they are limited in describing the quantitative impact of each factor on the occurrence of accidents. In order to make up for the deficiency of qualitative analysis, fault tree analysis and Bayesian network are widely used in the quantitative analysis of oil and gas fires. The combination of these two methods can give full play to their respective advantages and significantly improve the comprehensiveness, accuracy and practicability of risk analysis. Fault tree analysis [5] provides a qualitative analysis of the system fault structure, and reveals the logical relationship between various fault events through the hierarchical decomposition of the system fault modes. On the other hand, Bayesian network [6] can carry out probabilistic quantitative analysis on this basis, and transform qualitative reasoning into quantitative probabilistic inference, so as to evaluate the risk of the system more comprehensively. The combination of this method can not only effectively quantify the impact of various factors on the occurrence of accidents, but also reflect the change of system status in real time by dynamically updating the probability distribution, which provides a solid scientific basis and decision support for the prevention and emergency management of oil and gas fires.

On this basis, some scholars put forward the method of fuzzy fault tree to solve the non-deterministic problem of event state except certain occurrence or certain non-occurrence [7-9]. The introduction of fuzzy set processing into fault tree and Bayes has the characteristics of more flexible modeling, improving fault tolerance rate and dealing with uncertainty, so as to better support the analysis and decision of the system. According to the literature reviewed so far, fuzzy fault tree analysis method is more commonly used in fault diagnosis and reliability analysis of systems or equipment [10-11], and there is no report on the research on the causes of oil and gas storage and transportation fire accidents.

Therefore, this paper used a combination of fuzzy fault tree and Bayesian network method. Fault tree was used to explain the logical relationship between various basic events and the top event caused by basic events. Bayesian network provides a probability formula from the occurrence of basic events to the occurrence of top events. The mapping relationship between the fault tree and the Bayesian network connects the two, supplemented by the fuzzy set theory of natural language transformation of fuzzy numbers and comprehensive evaluation calculation to obtain the probability of basic event occurrence. With the basic probability of event occurrence, the logical relationship between various events, and the Bayesian total probability formula, it is possible to derive and calculate the probability of oil and gas fire risk and backward deduce the main factors that cause risk events, in order to achieve risk prevention and provide theoretical support for the refinement, precision, and efficiency of control work.

## 2. Statistics of Causal Factors for Oil and Gas Fire Risk

According to the statistical analysis of typical oil and gas fire and explosion accidents that have occurred in the past few years [12-13], it is found that the vast majority of oil and gas fire and explosion events are related to 9 approximate causal factors [14-16]. The names and number of events related to each causal factor are shown in Table 1.

*Table 1. Names and number of events related to causal factors*

| Incident name | Number of incidents |
|---|---|
| Regulatory loopholes | 54 |
| Managerial negligence | 13 |
| Explosion of petroleum products | 3 |
| Oil spillage | 9 |
| Harmful gases released from oil and gas | 10 |
| Failure of protective system | 8 |
| Damage to protective equipment | 18 |
| Equipment damage | 44 |
| Thermal radiation | 61 |

## 3. Risk Assessment Methods for Oil and Gas Fires

Starting from the probability of the occurrence of 9 causal factors and their logical relationship, this paper systematically sorts out how the causal factors lead to the formation of oil and gas fire risk events from bottom to top. Fault tree models and Bayesian network models are built for oil and gas fire risk analysis and assessment.

### 3.1 Establishment of Fault Tree Model

Fault tree modeling is a graphical method used to analyze the mechanisms of system failures and accidents. The method takes the top event not occurring as the top, and models various possible fault causes leading to the top event not occurring through a tree structure. By using the logical relationship between the events of the fault tree, the path leading to the non-occurrence of the top event can be identified and the corresponding protection measures can be proposed.With the help of fault probability calculation, probability values are assigned to each basic event in the fault tree, and these probability values can be assigned to higher-level nodes through and gates or gate logic gates, and so on, the failure probability of the whole system can be calculated [17-18].This method provides an important theoretical basis and practical tool for system reliability analysis, and is of great significance for improving the safety and reliability of the system. The fault tree model consists of event nodes, logic gates, and top events. Among them, event nodes represent possible faults or events that may occur in the system; logic gates represent the logical relationships between events; the top event indicates the final system failure or accident. By analyzing the fault tree, potential fault causes and modes in the system can be identified; the reliability and safety of the system can be evaluated; corresponding prevention and control strategies can be formulated.

Through careful observation and contemplation of the 9 causal factors, it is found that there exist both independent AND or logical relationships between events and priority or sequential relationships between them. Therefore, these factors can be constructed as a fault tree model. "Oil and gas fire" is selected as the research object; fault tree analysis method is applied; the FFTA model for emergency response of oil and gas fire accidents is established. The top event is a fire accident in the oil and gas area (T). Oil and gas fire incidents can be connected by four intermediate events: management issues ($M_1$), oil issues ($M_2$), protection issues ($M_3$), and hazardous states ($M_4$) through OR logic gates. This means that as long as one of the four intermediate events occurs, the top event can definitely occur. The four intermediate events can be connected by 9 causal factors through the AND or gate. The fault tree of oil and gas fires is shown in Figure 1, and the correspondence between basic events and codes is shown in Table 2.

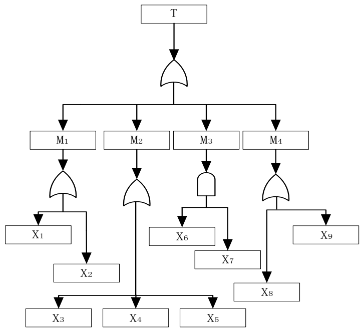


*Figure 1. Fault tree models*

*Table 2. Basic events*

| Code | Basic events |
|---|---|
| T | Oil and gas fires |
| M1 | Management issues |
| M2 | Oil problems |
| M3 | Failure of protection |
| M4 | Dangerous state |
| X1 | Regulatory loopholes |
| X2 | Managerial negligence |
| X3 | Explosion of petroleum products |
| X4 | Oil spillage |
| X5 | Harmful gases released from oil and gas |
| X6 | Failure of the protective system |
| X7 | Damage to protective equipment |
| X8 | Equipment damage |

### 3.2 Establishment of Bayesian Network Model

Bayesian network model is a graphical model based on Bayesian probability model, which is used to model the dependency relationship between random variables [18-19]. Bayesian networks use directed acyclic graphs to represent dependencies between variables, where nodes represent random variables and arrows point to represent dependencies between variables. Bayesian network theory is widely used in security engineering, financial risk, monitoring and detection, medical diagnosis and other fields. Bayesian network can be used to assess the risk of a project in security engineering, analyze the impact of various possible factors on the project risk, and establish the dependency relationship of risk factors for quantitative analysis and prediction. In medical diagnosis, Bayes can be used to monitor diseases, help doctors determine the source of diseases, verify the health status of patients, and provide accurate prediction for subsequent diagnosis by establishing the relationship between diseases. In the financial field, Bayesian network can be used for credit risk assessment, market risk assessment and money laundering risk identification, etc. By establishing the relationship between variables, including transaction amount, fund liquidity and historical repayment records, it can help maintain the security and stability of the financial system. Therefore, Bayesian networks provide powerful tools and methods for the modeling and analysis of complex systems, and provide important theoretical support for the processing of complex model reasoning decisions.A Bayesian network is a directed acyclic graph, where nodes represent random variables and arcs represent conditional probability dependencies between them. The “directed” refers to the fact that the direction of the arrows used to connect different nodes is fixed, and the order of the starting point and the ending point cannot be switched, which means that the logical relationship from cause to effect cannot be reversed. “Acyclic” refers to starting from any node and not being able to return to that point again through several edges, that is, there is no loop in the graph. The starting node connecting two non conditional independent random variable nodes is “cause”, and the ending node is “result”. The structure of a directed graph can simulate the causal structure of a modeling domain. When the structure is causal, the interactions between its variables provide visual information and can predict the effects of external manipulation. Bayes can perform forward inference to obtain the prior probability of child node events, and can also perform backward diagnosis to obtain the posterior probability of parent node events. Among them, the total probability formula is used to calculate the prior probability, and the Bayesian formula is used to calculate the posterior probability.

The topology of a Bayesian network is a Directed Acyclic Graph (DAG) where the nodes represent random variables and the arcs between the nodes represent conditional probability dependencies between these variables. "Directed" means that the edges used to connect different nodes have a fixed direction, that is, the order of the starting point and end point of the edge cannot be switched, which indicates that the direction of causality is irreversible, and thus reflects the logical flow from cause to effect. "No loop" means that there is no loop in the graph, that is, starting from any node, through a number of edges cannot return to the node. This acyclic nature ensures that there are no cyclic dependencies in the network, allowing the inference process to proceed unambiguously. In a Bayes network, the start node of a connection between two non-conditionally independent random variable nodes usually represents "cause" and the end node represents "effect". The structure of the directed graph can effectively simulate the causal structure in the modeling domain. When the structure of the network reflects causality in the real world, it is able to provide visual information on the interactions between variables. This structure not only helps to understand the complex dependencies within the system, but can also be used to predict the effects of external manipulation, and by manipulating one variable in the network, infer its effects on other related variables. Bayesian networks can perform both forward inference and backward diagnosis. Forward inference is used to calculate the prior probability of the child node event, that is, to infer the probability of the child node occurrence when the state of the parent node is known. This reasoning process relies on the total probability formula to obtain the total prior probability of the child node by integrating the probability information on different paths. Backward diagnosis is used to calculate the posterior probability of the parent node event, that is, to infer the occurrence probability of the parent node when the state of the child node is known. The posterior probability of the parent node is derived by updating the conditional probability distribution using Bayesian formula. Bayesian networks have significant advantages in modeling uncertain and complex systems. It can not only express the conditional dependence relationship between variables intuitively, but also provide quantitative analysis tools through accurate mathematical reasoning. Both in theoretical research and practical application, Bayesian networks show strong descriptive and predictive ability, and are widely used in medical diagnosis, risk assessment, decision support and other fields. Its flexibility and accuracy make it an important tool for dealing with complex data and uncertainty problems.

Assuming that the causal factors of a hazardous event are B1, B1,..., Bn, and that each causal factor is incompatible with each other, the probability of each causal factor occurring is greater than 0. The total probability formula is shown in Formula 1:

$$P(A) = \sum_{i=1}^{n} P[B_i]P[A|B_i] \qquad (1)$$

### 3.3 Mapping Relationship between Fault Tree and Bayesian Network

You Bingyu [20] proposed an innovative method combining fault trees and Bayesian networks. The method models the relationship between events through Bayesian networks, while the failure tree is utilized to model each failure basic event. Modeling using the combination of these two approaches provides a more comprehensive and accurate framework for system feasibility and stability analysis, which helps in system failure prediction and risk assessment. The use of fault trees and Bayesian networks at the same time makes the insight deeper, helping engineers better understand the operating mechanism and potential risks of the system, providing the optimal design path and decision support for the system.This research is important for the development of the safety engineering field.[21] The mapping relationship between Bayesian network and fault graph includes both image mapping and numerical mapping. Image mapping is to map the basic event Xi, intermediate event Mi, and top event T of the fault tree to the root, intermediate, and leaf nodes of the Bayesian network model, respectively. Numerical mapping maps the failure probability of the fault tree to a prior probability of the Bayesian network, while transforming the logic gates into a conditional probability table of the Bayesian network. The mapping principle of mapping fault trees to Bayesian networks is shown in Figure 2, and the logical mapping principle is shown in Figure 3.

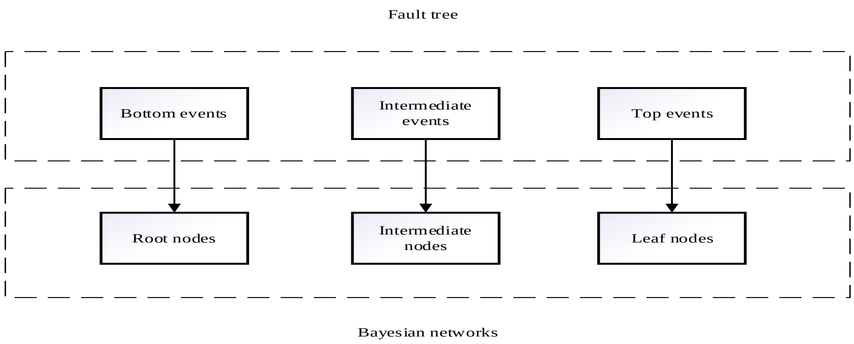


*Figure 2. Mapping rules*

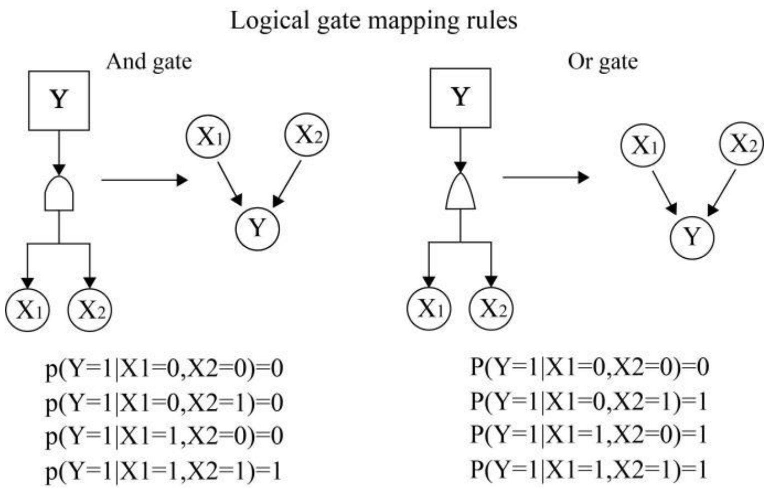


*Figure 3. Principle of logical mapping*

$X_1$ and $X_2$ are child nodes, and Y is the parent node. $P(Y=1|X_1=0,X_2=1)=1$ indicates that when the child node $X_1$ does not occur and $X_2$ occurs, the parent node Y occurs.

## 4. Forward Calculation of Probability of Occurrence

### 4.1 Fuzzy Set Theory

Fuzzy set theory is used to deal with fuzzy concepts and fuzzy information, and is a mathematical tool to deal with uncertainty. In the general concept of number theory, an element can only belong to the set or not belong to the set, and there are clear boundaries. However, in fuzzy set theory, there is no such strict subordination or non-subordination relationship, and only a membership degree is used to represent the dependency between elements and sets. After the membership degree is introduced, all kinds of fuzzy information can be better processed, thus supporting all kinds of fuzzy reasoning and decision making methods.In fuzzy set theory [22-24], an arbitrary domain U is given, and the independent variable x is any number in the domain U. If x corresponds to F(x) one by one under the correspondence rule F, F(x) is called the fuzzy set of U. The expression of fuzzy number formula F(j,k,l,m) is shown in Equation 2.

$$F(j,k,l,m)=\begin{cases}\frac{x-j}{k-j} & j\le x<k\\ 1 & k\le x<l\\ \frac{m-x}{m-l} & l\le x\le m\\ 0 & x<j, x>m\end{cases} \tag{2}$$

### 4.2 Expert Opinion Handling

In order to obtain the probability of 9 bottom events (root nodes) occurring, 15 oil and gas fire experts are invited to evaluate the occurrence (True) and non occurrence (False) states of 21 low events (root nodes) in oil and gas fires using a seven level natural language (seven level natural language and fuzzy tree are shown in Table 3). Due to differences and inaccuracies in natural language among experts, in order to eliminate the impact of these properties on the results, this paper uses the similarity fusion [25-27] processing method to perform comprehensive operations and reduce the differences in expert evaluations. The similarity fusion method is a commonly used technique in the fields of data mining and information retrieval, used to merge multiple similarity measures into a comprehensive similarity value. This technology can be used in a variety of neighborhood applications, including information recognition, book retrieval and autonomous driving. In many applications, it is necessary to compare the similarity between two objects, such as text similarity, image similarity, user interest similarity, etc. The similarity fusion method aims to integrate these similarity measures from different sources into a comprehensive similarity value, in order to more accurately represent the degree of similarity between objects. The advantage of similarity fusion method is to consider information from many aspects and multiple dimensions, and combine them organically according to different characteristics and similarity, so as to grasp and understand the relationship between data more comprehensively. Expert $Z_i$(i=1,2,…,15) evaluates a certain background event of an oil and gas fire using natural language. Based on the fuzzy numbers corresponding to natural language in Table 3, a series of calculations can be performed on the fuzzy numbers to verify the consistency and credibility of the expert language. The calculation process is as follows.

(1) Evaluation of similarity ($S_d$)

Evaluation of similarity refers to the degree of consistency between the evaluations of two experts on the same event. The higher the rating similarity, the more credible the evaluation results are. The evaluation of similarity $S_d$ is generally between 0 and 1 (0 indicates no similarity between the two, and 1 indicates almost identical). The evaluation formula of similarity $S_d$ for the evaluation of the same event by two experts m and n (m and n can be taken as 1, 2, 3,..., 15) is shown in Formula 3. In Formula 3, $q^i$ refers to the fuzzy number evaluated by expert q on node $X_i$:

$$S_d(Z_{m,n}) = 1 - \frac{1}{4}\sqrt{\sum_{i=1}^{4}(p^i - q^i)^2} \tag{3}$$

(2) Evaluation of average agreement ($A_{ge}$)

$$A_{ge}(Z_P) = \frac{1}{m-1}\sum_{\substack{q=1 \\ p\neq q}}^{M} S(R_q, R_q) \tag{4}$$

(3) Evaluation of relative agreement ($A_{ve}$)

$$A_{ve}(Z_P) = \frac{A_{ge}(Z_p)}{\sum_{p=1}^{M} A_{ge}(z_P)} \tag{5}$$

(4) Prediction of consensus coefficient ($C_{us}$)

$$C_{us}(Z_p) = \beta\omega(Z_p) + (1-\beta)A_{ve}(Z_p) \tag{6}$$

The weight of expert p is represented by ω(Zp), which is very important in similarity fusion method. The weight represents the contribution of the expert to the event. Although in reality, the weights of each expert should be different, in order to simplify calculations, experts with similar backgrounds are often selected to have the same weights. The relaxation factor β is a correction term.

(5) Comprehensive evaluation of fuzzy number $\mathrm{R_{i-G}}$

$$\mathrm{R_{i-G} = C_{us}(Z_1)R_1 + C_{us}(Z_2)R_2 + \ldots + C_{us}(Z_m)R_m} \tag{7}$$

(6) Comprehensive evaluation of fuzzy number $\mathrm{R_{i-G}}$ for deblurring $u_{i-G}$

$$u_{i-G} = \frac{\int F(x)xdx}{\int F(x)dx} \tag{8}$$

(7) $u_{i-G}$ is normalized, and the probability $\mathrm{P(X_{i-G})}$ for node Xi to have a state G∈(Ture or False) is:

$$\mathrm{P(X_{i-G})} = \frac{\mathrm{u_{i-G}}}{\sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{u_{i-j}}} \tag{9}$$

In Formula 9, n represents the total number of states at risk of system operation, and j represents the state.

*Table 3. Natural language changes and their fuzzy sets*

| Natural language variables | Fuzzy sets |
|---|---|
| Very Low(VL) | (0,0.1,0.1,0.2) |
| Low(L) | (0.1,0.2,0.2,0.3) |
| Fairly Low(FL) | (0.2,0.3,0.4,0.5) |
| Medium(M) | (0.4,0.5,0.5,0.6) |
| Fairly High(FH) | (0.5,0.6,0.7,0.8) |
| High(H) | (0.7,0.8,0.8,0.9) |
| Very High(VH) | (0.8,0.9,1.0,1.0) |

## 4.3 Root Node Occurrence Probability Calculation

This paper takes the risk event oil explosion ($X_3$) as an example, and calculates the probability $P(X_3)$ when the oil explosion state is true based on Formulas 3 to 9. The evaluation of 15 oil and gas fire risk experts on oil explosion ($X_3$) is shown in Table 4.

(1) Consistency test and comprehensive calculation of expert fuzzy evaluation

According to Formulas 3 to 6, the similarity $S_d$, average agreement $A_{ge}$, relative agreement $A_{ve}$, and consensus coefficient $C_{us}$ of risk event $X_3$ are calculated and tested separately. In order to ensure the objectivity and credibility of the evaluation results, this paper specifically selects 15 oil and gas risk experts with similar professional abilities and experience for evaluation. The weights of each oil and gas risk expert are equal, with ω=0.072, and the relaxation factor β is set to 0.5. According to the calculation results (as shown in Table 5), it is found that the similarity ($S_d$) of risk event $X_3$ for each oil and gas risk expert is as high as 0.9, indicating a high degree of consistency among oil and gas risk experts in evaluating the same risk event. In addition, the relative agreement ($A_{ve}$) remains largely consistent. Overall, these data further enhance the relative objectivity and credibility of the expert evaluation results of oil and gas risks, providing strong support for the analysis of this paper.

*Table 4. Evaluation results of experts on oil explosion incidents*

| Risk factors | Node status | Z1 | Z2 |
|---|---|---|---|
| X3(Explosion of petroleum products) | TRUE | FL | FL |
| | FALSE | H | FH |
| Risk factors | Node status | Z3 | Z4 |
| X3(Explosion of petroleum products) | TRUE | L | FL |
| | FALSE | FH | VH |
| Risk factors | Node status | Z5 | Z6 |
| X3(Explosion of petroleum products) | TRUE | FL | FL |
| | FALSE | FH | FH |
| Risk factors | Node status | Z7 | Z8 |
| X3(Explosion of petroleum products) | TRUE | L | FL |
| | FALSE | H | H |
| Risk factors | Node status | Z9 | Z10 |
| X3(Explosion of petroleum products) | TRUE | FL | FL |
| | FALSE | H | VH |
| Risk factors | Node status | Z11 | Z12 |
| X3(Explosion of petroleum products) | TRUE | FL | FL |
| | FALSE | H | H |
| Risk factors | Node status | Z13 | Z14 |
| X3(Explosion of petroleum products) | TRUE | FL | FL |
| | FALSE | H | H |
| Risk factors | Node status | Z15 | |
| X3(Explosion of petroleum products) | TRUE | FL | |
| | FALSE | H | |

*Table 5. Expert calculations of Sd, $A_{ge}$, $A_{ve}$, and $C_{us}$ for oil explosion events*

| Evaluation indicators | Metric values | Evaluation indicators | Metric values | Evaluation indicators | Metric values | Evaluation indicators | Metric values |
|---|---|---|---|---|---|---|---|
| Sd(Z12) | 0.956 | Age(Z1) | 0.977 | Ave(Z1) | 0.2513 | Cus(Z1) | 0.244 |
| Sd(Z13) | 0.933 | Age(Z2) | 0.971 | Ave(Z2) | 0.2492 | Cus(Z2) | 0.239 |
| Sd(Z14) | 0.966 | Age(Z3) | 0.976 | Ave(Z3) | 0.2444 | Cus(Z3) | 0.25 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |
| Sd(Z115) | 0.994 | Age(Z13) | 0.969 | Ave(Z13) | 0.2555 | Cus(Z13) | 0.246 |
| Sd(Z23) | 1 | Age(Z14) | 0.964 | Ave(Z14) | 0.2513 | Cus(Z14) | 0.228 |
| Sd(Z24) | 1 | Age(Z15) | 0.943 | Ave(Z15) | 0.2349 | Cus(Z15) | 0.239 |
| ... | ... | | | | | | |
| ... | ... | | | | | | |
| ... | ... | | | | | | |
| Sd(Z1314) | 0.964 | | | | | | |
| Sd(Z1315) | 1 | | | | | | |
| Sd(Z1415) | 0.899 | | | | | | |
| $R_{3\text{-True}}$ | 0.13365 | | 0.23367 | | 0.31342 | | 0.41344 |

(2) Probability $P(X_3)$ of occurrence of risk event $X_3$

According to Formula 8, the result $u_{i-G}$ of deblurring the comprehensive evaluation fuzzy number $R_{i\text{-}G}$ can be calculated, as shown in Formula 10.

$$u_{3-True} = \frac{\left[\int_{0.13365}^{0.23367}\left(\frac{x}{0.10002}\right)xdx + \int_{0.22537}^{0.31342}xdx + \int_{0.31342}^{0.41344}\left(\frac{0.41334-x}{0.10002}\right)xdx\right]}{\left[\int_{0.13365}^{0.23367}\left(\frac{x}{0.10002}\right)xdx + \int_{0.23367}^{0.31342}dx + \int_{0.31342}^{0.41344}\left(\frac{0.41334-x}{0.10002}\right)x\right]}$$

$$=0.2337 \tag{10}$$

Finally, according to Formula 9, the probability $P(X_3)$ of node $X_3$ occurring can be obtained, as shown in Formula 11.

$$P(X_3) = \frac{u_{3-True}}{u_{3-True+}u_{3-False}} = 0.329 \tag{11}$$

Referring to the above process, the probability of occurrence of underlying events $X_1$-$X_9$ can be calculated sequentially, as shown in Table 6.

*Table 6. Probability of occurrence of bottom level events at each node of oil and gas fires*

| Risk nodes | Probability of occurrence |
|---|---|
| X1 Regulatory loopholes | 0.64 |
| X2 Managerial negligence | 0.75 |
| X3 Explosion of petroleum products | 0.67 |
| X4 Oil spillage | 0.77 |
| X5 Harmful gases released from oil and gas | 0.71 |
| X6 Failure of protective systems | 0.75 |
| X7 Damage to protective equipment | 0.85 |
| X8 Equipment damage | 0.88 |
| X9 Thermal radiation | 0.93 |

### 4.4 Probability Calculation of Intermediate Nodes and Leaf Nodes

Taking the intermediate node $M_1$ (management issue) as an example, the probability of occurrence of the intermediate node is calculated. The intermediate node $M_1$ is connected to its root nodes $X_1$ (regulatory vulnerability) and $X_2$ (negligence by management) through an OR logic gate (as shown in Figure 1). According to the mapping relationship, the conditional probability of intermediate node $M_1$ is:

$$P(M_1 = True|X_1 = True, X_2 = True) = 1 \quad (12)$$
$$P(M_1 = True|X_1 = True, X_2 = False) = 1 \quad (13)$$
$$P(M_1 = True|X_1 = False, X_2 = True) = 0 \quad (14)$$
$$P(M_1 = True|X_1 = False, X_2 = False) = 0 \quad (15)$$

The Bayesian network total probability Formula 1 is used to calculate the occurrence probability of intermediate node $M_1$:

$$P(M_1 = True) = P(X_1 = True)P(X_2 = True) + P(X_1 = True)P(X_2 = False) + P(X_1 = False)P(X_2 = True) = 0.12 \quad (16)$$

Due to the fact that all research events are binary events, the probability of $M_1$ not occurring is:

$$P(M_1 = False) = 1 - P(M_1 = True) = 0.88 \quad (17)$$

By substituting the probability of each bottom event into the same conditional probability calculation formula for calculation, the probability of occurrence for each intermediate node and leaf node can be obtained. The specific calculation results are detailed in Figure 4. The comprehensive probability of occurrence of oil and gas fire risk is 0.71. It is worth noting that among these events, the probability of intermediate event $M_4$ (that is, dangerous state) occurring is the highest, significantly higher than the probability of other events occurring. Therefore, to ensure safety, special attention needs to be paid and corresponding protective measures should be taken, such as strengthening fire prevention measures, improving the safety performance of equipment, conduct a comprehensive safety risk assessment on a regular basis, formulate a detailed and effective danger emergency plan and conduct regular emergency drills, strengthening on-site monitoring and personnel training, in order to reduce the possibility of dangerous states occurring, and timely responding to potential dangerous states to ensure the safety and stability of oil and gas production and operation.

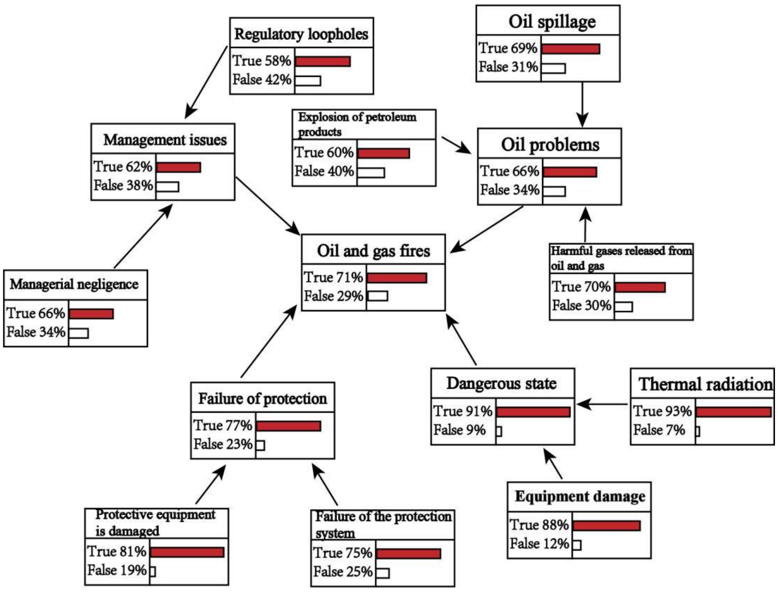


*Figure 4. Positive inference analysis of oil and gas fire risk*

## 5. Backward Inference

### 5.1 Backward Inference Calculation

Backward inference is the calculation of the posterior probability of occurrence for each node based on the assumption that the top event T (oil and gas fire) must be definitely occurred. The calculation formula is shown in Formula 18.

$$P(M_i = True|T = True) = \frac{P(T=True|M_i=True)P(M_i)}{\sum_{i=1}^{n} P(T=True|M_I=True)P(M_I)} \quad (18)$$

This paper is based on the assumption that the top event T (oil and gas fire) must be inevitably occurred, so the probability of setting its state to True is 100%. By using Formula 18 for backward inference, the posterior probabilities of each intermediate event are obtained. The specific results are shown in Figure 5. According to the analysis results in Figure 5, it is observed that the posterior probabilities of intermediate events $M_4$ (that is, dangerous state) and $M_3$ (that is, protection failure) are 0.91 and 0.8, respectively, which are significantly higher than the posterior probabilities of other intermediate events. It can be seen that in the process of oil and gas storage and transportation, it is necessary to attach great importance to the occurrence of these two risk events. Therefore, stricter safety measures should be taken to strengthen safety awareness training, including but not limited to providing fire emergency response training, on-site safety operation training, emergency evacuation drills, etc., to ensure that all staff can proficiently master the methods and skills to deal with dangerous situations. The safety operating procedures are strictly followed to minimize the probability of these two risk events and ensure the safety and stability of the oil and gas storage and transportation process.

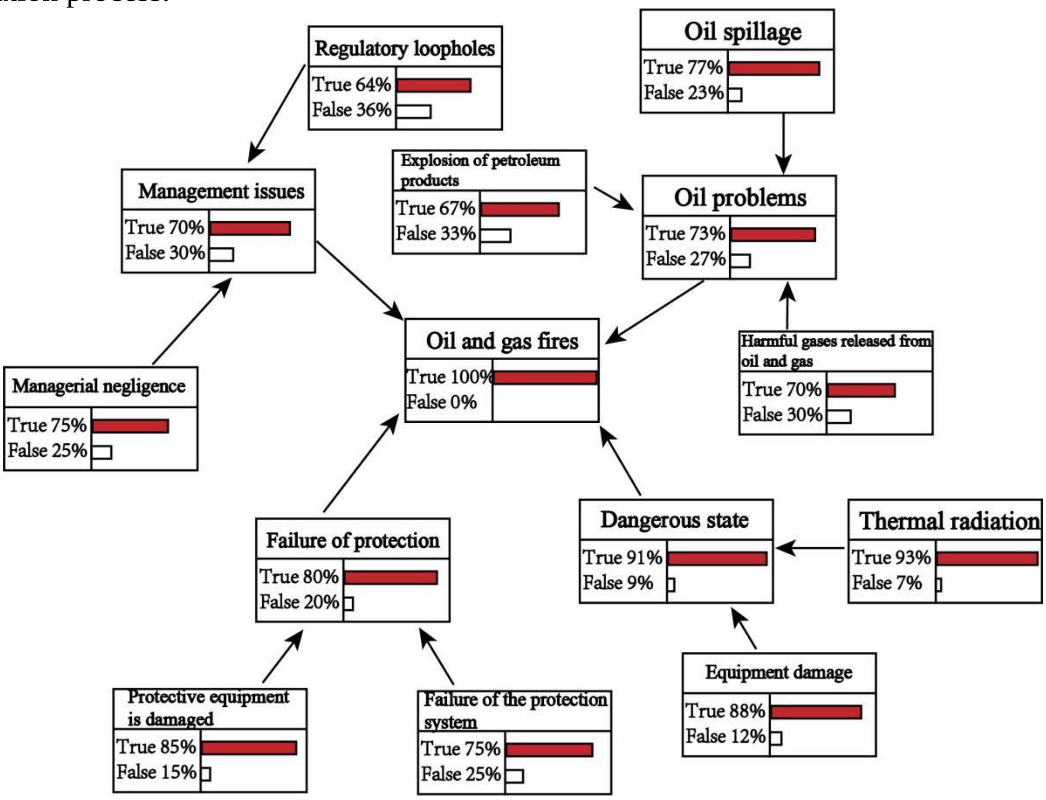


*Figure 5. Backward inference analysis of oil and gas fire risk*

## 5.2 Sensitivity

When there is a slight change in the probability of a bottom event occurring, the probability of an intermediate event also changes accordingly. However, different bottom events have different impacts on the change of the intermediate event. Sensitivity analysis[28-29] is used to calculate the probability of different bottom events occurring under the same intermediate event. The one with the highest probability of a bottom event occurring is the one that has the greatest impact on the intermediate event. This paper uses GENIE software for sensitivity analysis [30], and conducts sensitivity analysis on intermediate events $M_4$ (dangerous state) and $M_3$ (protection failure) with a high posterior probability of occurrence.

According to the results in Table 7, it can be seen that in the underlying events ($X_8$ equipment damage, $X_9$ thermal radiation) of intermediate event $M_4$ (hazardous state), $X_8$ equipment damage has the highest sensitivity, indicating that the impact of equipment damage on $M_4$ (hazardous state) is greater than that on $X_9$ thermal radiation. It can be inferred that the root cause of the dangerous state is equipment damage. Therefore, it is necessary to conduct a certain degree of pre job training for relevant personnel before work, standardize their operations, and strengthen the frequency of inspections by management personnel to reduce the probability of equipment damage.

According to the results in Table 8, it can be seen that the sensitivity of $X_6$ protective system failure is highest in the underlying events (X6 protective system failure, X7 protection equipment damage) of intermediate event $M_3$ (protection failure), indicating that the impact of protection system failure on $M_3$ (protection failure) is greater than that of $X_7$ protection equipment damage. It can be inferred that the root cause of protection failure is a malfunction of the protective system. Therefore, it is best to pre run and check whether the protective system is functioning properly before each operation of the system. It is necessary to strengthen the sense of responsibility of maintenance personnel and use information technology to detect faults in advance.

*Table 7. Sensitivity analysis of oil and gas fire risk M4*

| Risk nodes | Mutual information | Sensitivity % |
|---|---|---|
| Dangerous state | 0.68764 | 100 |
| The device is damaged | 0.64992 | 94.4 |
| Thermal radiation | 0.34766 | 50.6 |

*Table 8. Sensitivity analysis of oil and gas fire risk $M_3$*

| Risk nodes | Mutual information | Sensitivity % |
|---|---|---|
| Failure of protection | 0.59644 | 100 |
| Failure of the protection system | 0.54766 | 91.8 |
| Protective equipment is damaged | 0.26546 | 44.5 |

## 6. Conclusions

This paper proposed a model for predicting the risk of oil and gas fires. In order to maintain the objectivity of the evaluation results, this paper applied fuzzy set theory on the basis of the fault tree model. To calculate the prior probability of risk events, the fault tree model was mapped to a Bayesian network. The fault tree model provided clear and logical relationships between risk events, while Bayesian networks provided calculation formulas from root node probability to intermediate node probability and leaf node probability. The method in this paper can not only infer the prior probability of risk events forward, but also backward infer the posterior probability of intermediate events that trigger the top event and the fundamental event with the greatest impact. However, there are also some shortcomings. The basis for the statistical analysis of causal factors in this paper is based on previous accident reports. Over time, new causal factors may emerge and fail to be analyzed in a timely manner. Subsequent work can consider establishing a real-time updated system network and timely statistical analysis of new causal factors.

## Funding

1. This work was supported by Science and Technology Project of Sichuan Province (2023NSFSC1034)
2. This work was supported by the Fundamental Research Funds for the Central Universities (PHD2023-067)

## References

*[1]Fan Y. Management and maintenance measures of oil and Gas storage and transportation equipment [J]. Chemical Design Communication, 2021,47(11):7-8.*
*[2]Zhang Xinlin, Wei Ran. Analysis on Daily management and maintenance of oil and gas storage and transportation equipment in the Information Age [J]. China Petroleum and Chemical Standards and Quality, 21,41(04):109-111.*
*[3]Hosseinnia Davatgar B, Paltrinieri N, Bubbico R. Safety barrier management: risk-based approach for the oil and gas sector[J].Journal of Marine Science and Engineering, 2021, 9(7): 722.*
*[4]Dimaio F, Scapinello O, Zio E, et al. Accounting for safety barriers degradation in the risk assessment of oil and gas systems by multistate Bayesian networks[J]. Reliability Engineering & System Safety, 2021, 216: 107943.*
*[5]Zhang huan. The fault tree analysis method in the application of a certain commercial air source system troubleshooting [J]. Journal of internal combustion engine and accessories, 2023, (18) : 4-7. DOI: 10.19475 / j.carol carroll nki issn1674-957 - x. 2023.18.007.*
*[6]Zhang Weidong, Yu Youmin. Reliability Analysis of Subway Vehicle Air Conditioning System based on Bayesian Network [J]. Refrigeration and Air Conditioning,2023,23(08):78-83.*
*[7]Guo Zan. Risk Assessment of Cleaning device failure condition based on Fuzzy fault tree analysis [J]. Green science and technology, 2023, 25 (22) : 97-101. The DOI: 10.16663 / j.carol carroll nki LSKJ. 2023.22.040.*
*[8]Chen Yongjun, Li Xiaojian, Zhang Li, et al. Application of fault tree and Fuzzy Bayesian network in risk assessment of pipeline corridor operation and maintenance [J]. Journal of safety and environment, 2024, 24 (3) : 857-866. The DOI: 10.13637 / j.i SSN. 1009-6094.2023.0292.*
*[9]Sun Kuan, ZHANG Shuguang, Wang Zhu. Time and function based on fuzzy sets is harmful to the continuous changes in system analysis [J/OL]. Aviation engineering progress, 1-8 [2024-04-30]. http://192.144.236.110:8085/kcms/detail/61.1479.V.20240325.1316.006. html.*
*[10]Li Boyu, Liu Qigang. Research on safety risk factor decision of high-speed railway station based on fusion of fuzzy fault tree and Bayesian network [J/OL]. Railway standard design, 1-10 [2024-04-30]. https://doi.org/10.13238/j.issn.1004-2954.202212190002.*
*[11]Guo Jiming, Qi Jinping, Li Xingyun. Reliability Analysis of EMU Braking System Based on Fuzzy Dynamic Fault Tree[J]. China Mechanical Engineering, 2019, 30(13): 1585.*
*[12]Dong Tao, ZHAO Lei. Discussion on Safety Management of Oil Depot [J]. Science and Technology Innovation Review,2019,27(16):167-169.*

*[13]Wen Yacheng, KANG Yang. Safety Accident Analysis of Airport Oil Depot Based on Fault Tree [J]. Petrochemical Safety and Environmental Protection Technology,2021.*

*[14]Xu H J. Safety risk and Accident prevention and control of oil depot [J]. Chemical Industry Management, 2020.*

*[15]Guan Rui, TANG Huang. Analysis of Common Accidents in oil depot and Countermeasures [J]. Petrochemical Technology, 2021.*

*[16]Li T T. A brief discussion on the influencing factors and cognition of the safe operation of oil depots [J]. Chemical Management, 2019, 11.*

*[17]Yazdi M, Mohammadpour J, Li H, et al. Fault tree analysis improvements: A bibliometric analysis and literature review[J]. Quality and Reliability Engineering International, 2023, 39(5): 1639-1659.*

*[18]Kitson N K, Constantinou A C, Guo Z, et al. A survey of Bayesian Network structure learning[J]. Artificial Intelligence Review, 2023, 56(8): 8721-8814.*

*[19]Sevinc V, Kucuk O, Goltas M. A Bayesian network model for prediction and analysis of possible forest fire causes[J]. Forest Ecology and Management, 2020, 457: 117723.*

*[20]You BingyuLian Fumian, Meng Xianghai. Analysis of the causes of highway accidents in mountainous areas based on fault tree Bayesian nets[J]. Journal of Transportation Information and Safety, 2019, 37(4): 44-51.*

*[21]Wang C, Wang L, Chen H, et al. Fault diagnosis of train network control management system based on dynamic fault tree and Bayesian network[J]. IEEE Access, 2020, 9: 2618-2632.*

*[22]Lin S S, Shen S L, Zhou A, et al. Risk assessment and management of excavation system based on fuzzy set theory and machine learning methods[J]. Automation in Construction, 2021, 122: 103490.*

*[23]Pappas I O, Woodside A G. Fuzzy-set Qualitative Comparative Analysis (fsQCA): Guidelines for research practice in Information Systems and marketing[J]. International Journal of Information Management, 2021, 58: 102310.*

*[24]Liu Y, Jiang W. A new distance measure of interval-valued intuitionistic fuzzy sets and its application in decision making[J]. Soft Computing, 2020, 24: 6987-7003.*

*[25]Li Z H, Chen J, Bi J. Multiple object tracking with appearance feature prediction and similarity fusion[J]. IEEE Access, 2023.*

*[26]Rosso-Mateus A, Montes-y-Gómez M, Rosso P, et al. Deep fusion of multiple term-similarity measures for biomedical passage retrieval[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(2): 2239-2248.*

*[27]Sun J, Ding Y, Zhu X, et al. Extended Gaussian sphere and similarity fusion method for reassembly of 3D cultural relics[J]. Multimedia Tools and Applications, 2020, 79: 30187-30203.*

*[28]Razavi S, Jakeman A, Saltelli A, et al. The future of sensitivity analysis: an essential discipline for systems modeling and policy support[J]. Environmental Modelling & Software, 2021, 137: 104954.*

*[29]Liu B, Meng K, Dong Z Y, et al. Load balancing in low-voltage distribution network via phase reconfiguration: An efficient sensitivity-based approach[J]. IEEE Transactions on Power Delivery, 2020, 36(4): 2174-2185.*

*[30]Agbo C, Mehrpouyan H. Achieving Cyber-Informed Engineering Through Bayesian Belief Network and Sensitivity Analysis[C]//2023 10th International Conference on Dependable Systems and Their Applications (DSA). IEEE, 2023: 260-271.*