ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20240824001939
Original scientific paper

A Bayesian Network-Based Risk Assessment Model for Gas Pipeline Intelligent Management Systems

ChunXiao MEI*, JianXin TAN, Hao LI, JingTao CHANG, YiFan WANG, YiWei LU

Abstract: To enhance the accuracy and efficiency of gas pipeline risk assessment within intelligent management systems, this study proposes a Bayesian network-based data modeling and risk assessment framework. Traditional risk assessment methods rely heavily on statistical analysis and expert judgment, often struggling with uncertainty and interdependencies between risk factors. In contrast, Bayesian networks effectively model complex probabilistic relationships, providing a more dynamic and adaptive risk evaluation approach. This study integrates a cloud-based uncertainty processing model with Bayesian inference to improve risk prediction. Experimental validation using real-world gas pipeline monitoring data demonstrates that the proposed method achieves a 98.8% accuracy rate, significantly outperforming conventional techniques. Additionally, the assessment period is reduced to approximately two months, enhancing real-time decision-making capabilities. These findings highlight the potential of Bayesian networks to transform gas pipeline safety management by offering precise, efficient, and adaptive risk assessment models. Future research will explore integrating real-time IoT sensor networks and machine learning-based anomaly detection to further optimize predictive capabilities.

Keywords: accident prediction; bayesian networks; gas pipeline; intelligent management; risk assessment

1 INTRODUCTION

The intelligent management system for gas pipelines is a system designed for the intelligent monitoring and management of urban gas pipelines, which plays a critical role in ensuring the safety of residents' lives and urban operations [1, 2]. The core advantage of the smart gas management system lies in its distributed monitoring network, which is based on Internet technology [3, 4]. Beyond monitoring, the smart gas management system also includes warning and emergency response capabilities. The system can monitor real-time data such as operation status, temperature, pressure, and other parameters of gas pipelines. Once abnormal data is detected, the system immediately issues a warning to notify relevant personnel to take appropriate measures [5, 6]. Traditional risk assessment methods typically rely on statistical and engineering experience analysis, which makes it difficult to fully capture the potential risks of gas pipelines in complex environments. This limitation can affect the accuracy and reliability of risk assessments [7, 8]. Bayesian networks offer more reliable evaluation results through logical reasoning and probability calculations due to their ability to handle uncertainty in complex systems [9]. Bayesian networks have become a powerful tool for gas pipeline risk assessment because of their advantages in managing uncertainty and modeling complex systems. They can integrate prior knowledge with real-time monitoring data, dynamically updating the system's status and quantifying uncertain factors, allowing for more accurate risk assessments of gas pipeline failures. In addition, Bayesian networks are capable of modeling the complex dependencies between risk factors and can adaptively update the assessment as new data is acquired, further improving the accuracy of risk assessments. Based on this, we propose a risk assessment method for gas pipeline intelligent management systems that utilizes Bayesian networks and incorporates a cloud model to manage uncertainty and fuzziness within the system. The innovation of this research lies in the combination of Bayesian networks' ability to capture complex causal relationships and cloud models' capacity to process data uncertainty and ambiguity, providing a new framework for risk assessment in intelligent gas pipeline management systems. This approach not only enhances the accuracy and reliability of risk identification but also offers new ideas and technical support for intelligent gas pipeline safety management.

The research content is primarily divided into four parts. The first part provides an overview of the Intelligent Management System (IMS) for gas pipelines, along with data modeling and risk assessment based on Bayesian networks. The second part introduces the data modeling and risk assessment method based on cloud models and Bayesian networks within the IMS for gas pipelines. The third part validates the proposed model. The fourth part discusses the research results and outlines future prospects.

2 LITERATURE REVIEW

The risk assessment of gas pipelines is crucial for ensuring the safety of residents' lives and production. Several assessment methods are currently available. Regarding the question of whether natural gas pipelines can be economically used for hydrogen transportation, Cerniauskas et al. derived the cost function for the selected pipeline real location method. Their results indicated that, compared to building a new hydrogen pipeline system, the transportation system's cost was reduced by 30% [10]. Yuji et al. proposed a quantitative risk assessment method for natural gas pipeline network safety management based on the fuzzy analytic hierarchy process (AHP) and an improved coefficient of variation. Their findings showed that this method could effectively assess the risks of natural gas pipeline networks [11]. Wickramasinghe et al. explored the applicability of the Naive Bayes algorithm in different environments and discussed the advantages, disadvantages, and vulnerabilities of the algorithm. The experimental results demonstrated the effectiveness of their proposals [12]. Compared to the proposed method, the fuzzy AHP used by Yuji et al. still relies on expert evaluations when determining weights. The subjectivity in this process may affect the consistency of the final evaluation results. Additionally, insufficient consideration of the interrelationships between risks may not fully reflect the risk characteristics of complex systems. The Naive Bayes algorithm used by Wickramasinghe et al., on the other hand, may be limited by its assumption of independence when dealing with complex and multidimensional risk data, potentially affecting accuracy. Furthermore, this method lacks a dynamic response mechanism to changes in data, which means it may not adapt quickly to environmental changes. In contrast, the proposed method effectively captures and processes uncertainty and complexity in risk assessment using Bayesian networks. This approach more accurately reflects pipeline status and potential risks. Compared to traditional methods, the research provides a more comprehensive understanding of the mutual influences between different risk factors, offering decision-makers a broader perspective on risk assessment.

In summary, both domestic and foreign researchers have proposed various methods for pipeline risk assessment, achieving certain results. However, few scholars have utilized Bayesian networks to address these challenges. Therefore, Bayesian networks are introduced for the quantitative evaluation of risk assessment. This approach is expected to provide valuable insights and application benefits for hazard prevention in gas pipelines.

3 RESEARCH METHODOLOGY

Traditional risk assessment methods often rely on historical data and statistical models. Although they can provide useful analytical references, they frequently exhibit limitations when addressing complex accidents and changing environmental factors [13-15]. Cloud modeling is an emerging theory and technology that can effectively manage uncertainty and ambiguity. Its core idea overcomes many challenges faced by traditional methods in dealing with complex systems by converting qualitative expert evaluations into quantitative data. To address the gas safety issues caused by the expansion of urban gas pipeline networks, a semi-quantitative evaluation method is proposed. This method uses the cloud model to derive the weight matrix of gas evaluation indicators, and then iterates the expert comment matrix using the Delphi method. Finally, the method is presented in the form of a cloud map.

3.1 Gas Pipeline Risk Assessment Index System and Assessment Process

Before assessing gas pipelines in the IMS, an evaluation index system is established, focusing on the characteristics of gas pipelines. To better analyze these characteristics, a comparison is made between gas pipelines and urban long-distance pipelines, as shown in Fig. 1. In Fig. 1, there are significant differences between gas pipelines and urban long-distance pipelines in terms of detection methods, laying methods, pipeline connection lines, and ground subsidence phenomena [16-18]. Based on these differences, and considering the evaluation methods for urban long-distance pipelines, along with an analysis of previous gas pipeline safety accidents, a gas pipeline evaluation index system is established.

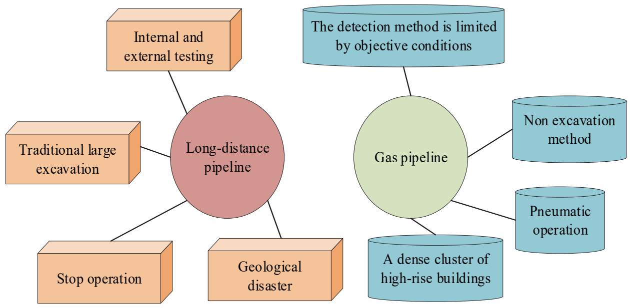

Figure 1 Comparison of characteristics between gas pipelines and urban long-distance pipelines

Building on this risk assessment index system, a gas pipeline risk assessment method based on the Delphi method is proposed. The Delphi method systematically collects and integrates expert opinions through anonymous questionnaires and multiple rounds of feedback, reducing the inconsistency of results caused by the subjective biases of individual experts. This method leverages the rich experience of experts to provide reliable risk assessment data in the presence of sufficient historical data, thereby enhancing the scientific rigor of the assessment. Additionally, cloud models are used to convert fuzzy expert judgments into quantitative data. Cloud models are particularly suited for handling uncertainty and ambiguity, as risk assessments often involve fuzzy expert judgments. These models effectively quantify such fuzzy information [19-21]. The technical roadmap for this method is shown in Fig. 2.

In Fig. 2, the cloud model process begins with clarifying the evaluation objectives and indicators. Expert opinions on the evaluation indicators are collected through questionnaire surveys, and expert feedback is integrated using the Delphi method to obtain consistent evaluation data. Next, the qualitative assessments are converted into quantitative data, and key parameters of the cloud model are calculated. The weights of the evaluation index system are determined using cloud models and the Analytic Hierarchy Process (AHP). Following this, the Delphi method is applied to iterate the expert comment matrix, and the two matrices are combined to compute the final evaluation results. Based on statistical mathematics and fuzzy mathematics, the randomness and fuzziness between uncertain linguistic values and exact numerical values are consistently characterized, facilitating a natural transformation between qualitative linguistic values and quantitative numerical values. In the cloud model, the certainty of qualitative concepts is represented by Eq. (1).

$$\mu c(x): U \rightarrow [0, 1], \forall x \in U, x \rightarrow \mu c(x) \tag{1}$$

In Eq. (1), U represents a quantitative set composed of deterministic numbers. c represents the qualitative concept of quantitative sets. μ represents a certain implementation of the qualitative concept in quantitative concentration. x represents the quantitative value in the quantitative set. The expression form in the cloud model is shown in Eq. (2).

$$\mu c(x) = e^{-\frac{(x_i - Ex)^2}{2(En_i)^2}} \tag{2}$$

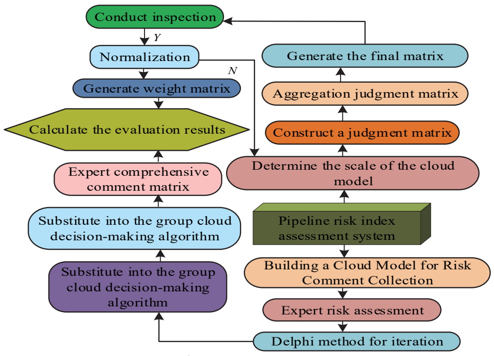

Figure 2 Technical roadmap for semi-quantitative evaluation results of gas pipelines

In Eq. (2), $\mu c(x)$ signifies the model membership degree. *Ex* represents the expected distribution of cloud droplets in the domain, which can represent qualitative concepts and define the position of the cloud in the domain, i.e., the center position of the cloud. *En* represents the entropy, corresponding to the variance in the likelihood distribution based on statistical theory. According to statistical knowledge, the range of *Ex* to 3*En* should cover 99% of conceptually acceptable elements. The first step in using the cloud model for weight calculation is to determine the model scale. The scale can be obtained using Eq. (3).

$$\frac{1}{c\left(Ex_{ij}, En_{ij}, He_{ij}\right)} = c\left(\frac{1}{Ex_{ij}}, \frac{En_{ij}}{\left(Ex_{ij}\right)^2}, \frac{He_{ij}}{\left(Ex_{ij}\right)^2}\right) \tag{3}$$

In Eq. (3), *He* represents the hyperentropic, which corresponds to the hyperparameter of statistical theory, specifically the uncertainty measure of entropy. It indirectly reflects the thickness of the cloud. The larger the hyperentropic, the greater the dispersion of cloud droplets. *ij* represents the importance of safety condition influencing factor *i* to safety condition influencing factor *j*. By combining the cloud model with AHP, the judgment matrix of expert comments is calculated. The distribution expectation calculation is shown in Eq. (4).

$$Ex_i = \prod_{j=1}^{n} Ex_{ij} \tag{4}$$

In Eq. (4), *n* signifies the order of the matrix. The entropy is shown in Eq. (5).

$$En_i = \prod_{j=1}^{n} Ex_{ij} \sqrt{\sum_{j=1}^{n} \left(\frac{En_{ij}}{Ex_{ij}}\right)^2} \tag{5}$$

The super entropy is shown in Eq. (6).

$$He_i = \prod_{j=1}^{n} Ex_{ij} \sqrt{\sum_{j=1}^{n} \left(\frac{He_{ij}}{Ex_{ij}}\right)^2} \tag{6}$$

It is normalized to obtain weights, and the weight matrix is then checked and judged as qualified. In addition to the weight matrix, the risk assessment results of gas pipelines also require an expert evaluation matrix. The study first establishes a cloud model of the risk comment set and iterates the given risk comments through the Delphi method. The Delphi method is a systematic programming approach. In this study, 10 experts from diverse backgrounds, including researchers from research institutions, technical experts in the industry, and management personnel, are selected to ensure the diversity and comprehensiveness of opinions. Expert opinions are collected using the Delphi method. In the initial stage, a detailed questionnaire covering various aspects related to gas pipeline risk assessment is developed. The questionnaire includes different risk factors and the determination of language variables, and experts fill out the questionnaire anonymously. The research team summarizes and analyzes expert feedback through multiple rounds of feedback. Each round of the questionnaire focuses on areas closely related to the feedback from the previous round, gradually reducing uncertainty. In the final round of feedback, the final evaluation results are presented to all participating experts, and their opinions are solicited. After repeated consultation, induction, and modification, the results are finally summarized into a consensus among experts as the prediction result [22-24]. The specific flowchart of Delphi method is shown in Fig. 3.

In Fig. 3, the first step is to identify the problem to be solved. The investigator drafts a survey form and consults experts through letters following established procedures. Experts then submit their opinions anonymously. After several rounds of consultation and feedback, the opinions gradually become more centralized.

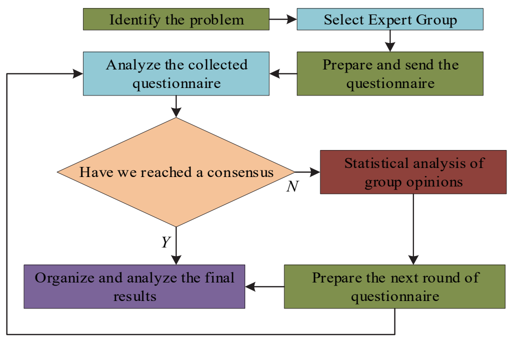

Figure 3 Delphi process flowchart

Finally, a highly accurate collective judgment result is obtained. The expert decision results are integrated, as shown in Eq. (7).

$$Ex_i = \frac{\left(Ex_i\right)_1 + \left(Ex_i\right)_2 + \ldots + \left(Ex_i\right)_N}{N} \tag{7}$$

In Eq. (7), *N* signifies the expert involved in the decision-making process. The entropy calculation is shown in Eq. (8).

$$\begin{aligned} En_i = \frac{1}{6}\Big[& \max_{I}\left\{\left(Ex_i\right)_I + 3\left(En_i\right)_I\right\} - \\ & - \min_{J}\left\{\left(Ex_i\right)_J - 3\left(En_i\right)_J\right\}\Big] \end{aligned} \tag{8}$$

The super entropy is shown in Eq. (9).

$$He_i = \frac{\left(He_i\right)_1 + \left(He_i\right)_2 + \ldots + \left(He_i\right)_N}{N} \tag{9}$$

The final comment set matrix is divided into 5 levels, with levels 1-5, where lower levels indicate higher risks. Finally, the weight matrix and comment matrix obtained in the previous steps are comprehensively calculated to yield the evaluation result for the gas pipeline.

3.2 Pipeline Risk Assessment and Incident Rating Evaluation Based on Bayesian Networks

The risk assessment method for gas pipelines based on cloud models has certain advantages in dealing with uncertain factors, but its implementation is relatively complex. Additionally, for developers, the understanding and application cycle can be quite long [25-27]. To make the evaluation system more convenient for use in the IMS of gas pipelines, a quantitative risk assessment method based on Bayesian networks has been developed. The implementation steps for this method are displayed in Fig. 4.

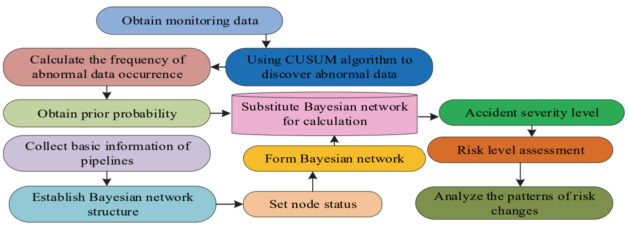

Figure 4 Implementation steps of a quantitative risk assessment method for gas pipelines based on Bayesian networks

In Fig. 4, the monitoring data is first collected from the monitoring system. Then, the Cumulative Sum (CUSUM) algorithm is used to calculate the frequency of change points and obtain the prior probability. Prior probability refers to the belief or estimation of the value of an event or variable before any data is observed. In Bayesian networks, prior probabilities are used to update these beliefs. When new data or evidence appears, the prior probability can be adjusted using the posterior probability formula to provide more accurate estimates. In Bayesian networks, the derivation of conditional probabilities is achieved by establishing dependency relationships between variables. For each node, its conditional probability table (CPT) is clearly defined. The CPT contains the probability distribution of each node given the state of its parent node. If node A is the parent node of node B, the value of $P(B \mid A)$ can be set through historical data analysis or expert judgment to calculate the probability of accidents occurring. The CUSUM algorithm detects change points in data sequences. However, it may increase the computational burden when processing large datasets or high-frequency monitoring data in real-time, especially when performing multiple recalculations. The CUSUM method is used to calculate the frequency of changes in various monitoring data to update the prior probabilities in the Bayesian network. The workflow of the CUSUM method involves collecting historical data on pressure, flow rate, and temperature from the pipeline monitoring system, followed by data preprocessing to fill in missing values and handle outliers. The mutation parameters are then determined, and the cumulative sum is calculated to construct a control chart, with upper and lower control limits set. When these control limits are exceeded, they are marked as mutation signals to detect abnormal conditions during pipeline operation. The time series X satisfies Eq. (10).

$$X = \left\{X\left(k\right)\right\} \quad k = 1,\, 2,\, 3,\, \ldots,\, n \tag{10}$$

In Eq. (10), $X(k)$ satisfies Eq. (11).

$$X\left(k\right) = \mu + Mk\left(k < \tau\right) + N_k\left(k \geq \tau\right) \tag{11}$$

In Eq. (11), μ represents the mean of X. M_k and N_k respectively represent different indicator functions. τ represents the time of mutation occurrence. For the unilateral CUSUM control chart, the relationship is given by Eq. (12).

$$\begin{cases} X_0 = 0 \\ X_k = \max\left\{0, X_{k-1} + S_k\right\} \end{cases} \tag{12}$$

In Eq. (12), S_k satisfies the relationship, as shown in Eq. (13).

$$S_k = X_k - \mu_0 - \beta \tag{13}$$

In Eq. (13), μ_0 represents the mean value when no mutation occurs, and β represents the standard deviation. A Bayesian network model is constructed based on the characteristics of the target to be evaluated while obtaining prior probabilities. A Bayesian network is a probabilistic graphical model used to express variables and their conditional dependencies. In gas pipeline risk assessment, nodes represent various risk factors or events, such as pipeline damage, external interference, equipment aging, accidents, etc. The edges represent the conditional dependency relationships between nodes. The Bayesian calculation is displayed in Eq. (14).

$$P\left(y \middle| x\right) = \frac{P\left(x \middle| y\right)P\left(y\right)}{P\left(x\right)} \tag{14}$$

In Eq. (14), $P(y \mid x)$ signifies the posterior probability. $P(x \mid y)$ signifies conditional probability. $P(y)$ signifies the prior probability of event y. $P(x)$ represents the prior probability of event x. A Bayesian network is a probabilistic graphical model that removes conditional probability independence. One property of a Bayesian network is the local Markov property. Given the parent node of a node in a directed acyclic graph, the node is independent of all its non-successor nodes [28-30]. For a Bayesian network B, there is a relationship shown in Eq. (15).

$$B = < B_G, B_P > \tag{15}$$

In Eq. (15), B_G represents its network structure. B_P is the conditional probability distribution in Bayesian networks. According to the different levels of accidents that occur, judgments can be made based on the principle of the highest level of accident severity [31, 32]. The level assessment of the accident occurrence is shown in Fig. 5.

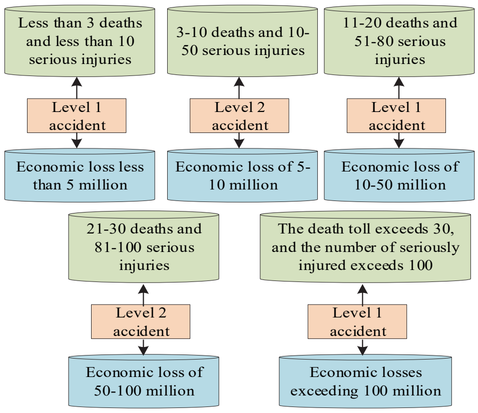

Figure 5 Accident level assessment

In Fig. 5, accidents with fewer than 3 casualties, fewer than 10 serious injuries, and economic losses below 5 million yuan are classified as level 5 accidents. The level 1 accident classification standard is that the number of fatalities exceeds 30, the number of seriously injured exceeds 100, or the economic losses exceed 100 million yuan. The risk assessment adopts the risk matrix method, which performs a comprehensive judgment based on the probability level and accident loss level, as shown in Fig. 6.

Risk level		Loss level				
		Catastrophic	Very serious	Serious	General	Negligible
Likelihood level	Frequent	I	I	I	II	III
	Possible	I	I	II	III	III
	Occasional	I	II	III	III	IV
	Rare	II	III	III	IV	IV
	Impossible	III	III	IV	IV	IV

Figure 6 Risk matrix table

In Fig. 6, the red area signifies level 1, the yellow area stands for level 2, the blue area signifies level 3, the purple area signifies level 4, and the green area signifies level 5. Under conditions of impossible probabilities and negligible loss levels, the accident level is higher, representing a lower accident intensity. Conversely, at levels of catastrophic loss and frequent likelihood, lower accident levels correspond to higher severity of accidents.

4 RESULTS AND DISCUSSION

To demonstrate the effectiveness of the data modeling and risk assessment method based on Bayesian networks, monitoring data from a gas company in a specific city for a certain section of a gas pipeline is selected. First, the risk assessment method based on cloud models is evaluated. Then, the quantitative risk assessment method for gas pipelines using Bayesian networks is applied to evaluate the unified pipeline section. Finally, a commonly used scenario model evaluation method is chosen as the comparison method to compare with the two proposed methods. The experimental implementation includes an Intel i5 processor, 6 GB RAM, 500 GB SSD, Windows 10 operating system, Python 3.7 programming environment, Matplotlib visualization library, and scikit-learn machine learning library. Real-world monitoring data is collected, which includes historical pressure, flow rate, and temperature data of gas pipelines. This data can come from sensors, control systems, or maintenance records. Missing values are handled using interpolation and median filling methods to ensure data continuity. Outliers are identified and handled using the Z-score statistical method to ensure data validity. Numerical features are standardized or normalized to ensure all features are on the same scale. The data is divided into a training set (70%) and a testing set (30%) to evaluate model performance. The research involves real-world monitoring data and adheres to the following ethical and data privacy principles: all data collection and use comply with relevant laws, regulations, and industry standards, and necessary licenses and approvals are obtained. All participants have been notified prior to using the data to ensure they understand the purpose of the data collection and its potential impact.

4.1 Risk Assessment and Analysis of Gas Pipelines Based on Cloud Models

To demonstrate the effectiveness of the cloud-based gas pipeline risk assessment method proposed in the research, the weight matrix and expert comment matrix are calculated separately based on the established evaluation index system. The comprehensive evaluation is then obtained. The experiment focuses on a section of gas pipeline located in the city center, which is 1037.55 meters long and surrounded by commercial and office areas. According to the risk indicator evaluation system, the weight cloud map for pipeline damage, risk emergency response, and accident severity, as well as the comprehensive evaluation cloud map of the three indicators, are shown in Fig. 7.

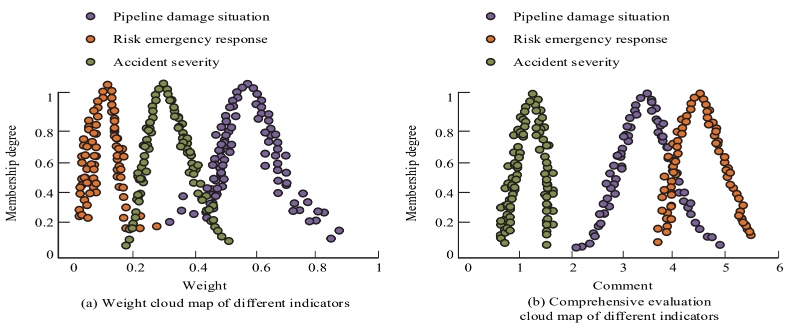

Figure 7 Cloud map of three indicator weights and comprehensive comments

In Fig. 7a, the weight distribution of each indicator in the evaluation system is relatively scattered. The overlapping area of pipeline damage at weight 0.6 is relatively large, and the distribution range of cloud droplets is moderate, with moderate cloud thickness. From the comparison of cloud maps, the overlapping area of risk emergency response at weight 0.1 is relatively large, and the distribution range of cloud droplets is small, with thick cloud layers. The overlapping area of green cloud droplets at weight 0.3 is relatively large, and the distribution range of cloud droplets is moderate, with moderate cloud thickness. In Fig. 7b, regarding the comprehensive comments of experts, the distribution of pipeline damage falls between 2-5, with the densest distribution around 3.6. The distribution of risk emergency response varies between 0.7-1.8, with the densest distribution around 1.2. The accident severity index has a distribution between 3.8-5.5, with the densest distribution around 4.5. Experimental data show that in the semi-quantitative gas pipeline risk assessment index system, the weight proportion of the accident severity index is the highest, while the weight proportion of pipeline damage is the lowest. The expert comments are (3.715, 0.406, 0.136), (1.348, 0.143, 0.073), and (4.535, 0.168, 0.067), respectively. The weight distribution of pipeline damage risk is around 0.6, indicating that this indicator occupies an extremely important position in the entire risk assessment. Pipeline damage is the main cause of gas leakage and other accidents. Therefore, when evaluating the safety of gas pipelines, the physical state and integrity of the pipeline should be given special attention. The low weight of risk emergency response capability suggests that, although its importance in improving safety and reducing accident losses cannot be ignored, its direct impact on the risk assessment is relatively limited in the current situation. The weight of accident severity is around 0.3, reflecting the importance of the potential consequences once an accident occurs. The evaluation results show good certainty and general acceptance. Based on the weights and comprehensive evaluation levels of the above indicators, the risk level evaluation of this section of the pipeline is shown in Fig. 8.

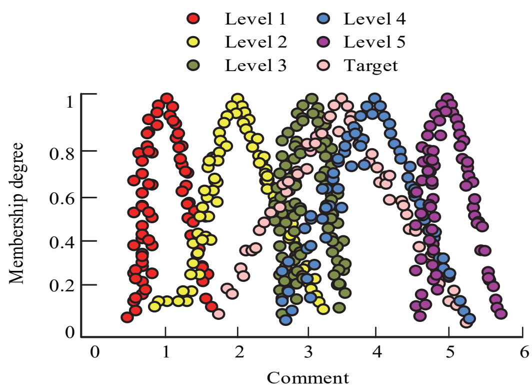

Figure 8 Cloud map of final evaluation level of pipeline

According to Fig. 8, the final gas pipeline risk assessment results yielded an evaluation level of (3.622, 0.415, 0.138). Cloud droplets representing red level 1 risk were distributed between 0.5 and 1.8. In the comparison of cloud maps, the distribution range of level 1 risk cloud droplets was moderate, with moderate cloud thickness. Level 2 risk yellow cloud droplets were distributed between 0.8 and 3.2, and the distribution range of level 2 risk cloud droplets was also moderate, with moderate cloud thickness. Level 3 risk was distributed between 2.8 and 3.5, with a smaller range of cloud droplets and thinner cloud thickness. The evaluation objects had a distribution between 1.9 and 5.2, with the densest distribution around 3.5. Level 4 was distributed between 2.8 and 5.2, and the distribution range of level 4 risk cloud droplets was relatively large, with moderate cloud thickness. Level 5 risk was distributed between 4.8 and 5.7, with a relatively small range of cloud droplets and thin cloud thickness. The experimental data showed that the final evaluation level, combined with the principle that the distribution expectation of the cloud model should cover 99% of conceptually acceptable elements within a range of three times entropy, indicates that this pipeline falls into level 4. The evaluation results show that the pipeline has been assessed as high-risk, suggesting that current operational and management measures have not effectively mitigated existing risks. This result has prompted management to immediately increase vigilance and focus on the safety monitoring and maintenance of the pipeline. Based on these evaluation results, it is necessary to implement stricter monitoring mechanisms in the daily management of pipelines, strengthen regular pipeline inspections, and introduce real-time monitoring technology to promptly identify and address potential problems.

4.2 Quantitative Risk Assessment Analysis of Gas Pipelines Ground on Bayesian Networks

To demonstrate the effectiveness of the quantitative risk assessment method based on Bayesian networks, the study first analyzes pressure and flow data anomalies that have occurred in the pipeline section over recent years, and calculates the risk probability. Subsequently, the risk probability is combined with the Bayesian network to obtain the risk level, and the evolution of the risk over time is analyzed. Finally, to evaluate the efficiency of risk assessment with a large volume of data, the commonly used scenario model evaluation method is selected to conduct comparative experiments on the three methods. The AHP diagram of flow data for pipelines at different time periods is shown in Fig. 9.

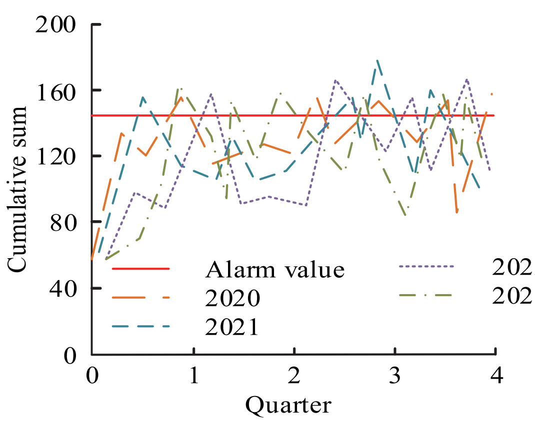

Figure 9 AHP diagram of pipeline flow data at different time periods

According to Fig. 9, in 2020, the traffic data exceeded the alarm value five times, indicating five abnormal situations in the traffic data. In 2021, there was one abnormal situation in the traffic data during the first to third month, no abnormal situations in the next three months, two abnormal situations in the third quarter, and one abnormal situation in the fourth quarter. The abnormal data in 2022 were concentrated in the second half of the year, with one occurrence in the first half and three in the second half. In 2023, there were six instances of abnormal traffic data, with one in the first and third quarters, and two in the second and fourth quarters. From the monitoring data, the pipeline monitoring system can easily and quickly gather data for each time period. Combined with AHP, the flow data for different time periods can be accurately monitored. The risk probability can be determined by the number of risk occurrences in the pipeline, as shown in Tab. 1.

Table1 Probability of pipeline risk occurrence at different time periods

	2020	2021	2022	2023
Number of abnormal pressure data occurrences	6	3	5	4
Probability of abnormal pressure data / %	0.6	0.25	0.49	0.37
Number of abnormal traffic data occurrences	5	4	4	6
Probability of abnormal traffic data / %	0.49	0.37	0.37	0.6

According to Tab. 1, in 2020, there were six instances of abnormal pressure data, with a probability of 0.6%. The frequency of abnormal traffic data occurrences was five times, and the probability of abnormal traffic data occurrence was 0.49%. In 2021, there were three instances of abnormal pressure data, with a probability of 0.25%. The frequency of abnormal traffic data occurrences was four times, with a probability of 0.37%. In 2022, there were five instances of abnormal pressure data, with a probability of 0.49%. The frequency of abnormal traffic data occurrences was four times, with a probability of 0.37%. In 2023, there were four instances of abnormal pressure data, with a probability of 0.37%. The frequency of abnormal traffic data occurrences was six, with a probability of 0.6%. According to monitoring data and prior probabilities from Bayes' theorem, the probability of pipeline pressure and flow risk can be accurately determined, demonstrating the accuracy of the Bayesian network quantitative evaluation method in risk probability prediction. Frequent abnormal events prompt decision-makers to pay timely attention to the operational status of pipelines and take necessary monitoring and maintenance measures. From the annual data changes, it is evident that the effectiveness of management strategies can directly influence the frequency of abnormal events. The significant decrease in the number of occurrences in 2021 may be attributed to strengthened maintenance and management measures for pipelines. Based on the abnormal increase in traffic in 2023, it is recommended to prepare emergency plans for potential risks in advance. To determine the pattern of risk occurrence and implement preventive measures, a statistical analysis is conducted on the numerical evolution from 2020 to 2024, as shown in Fig. 10.

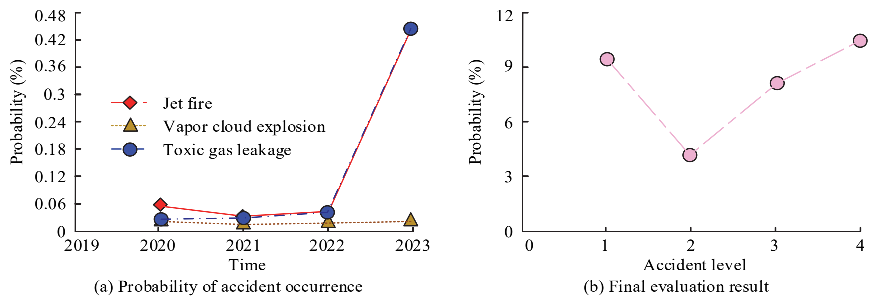

Figure 10 AHP diagram of pipeline flow data at different times

According to Fig. 10a, the probability of a jet fire accident occurring in 2020 was 0.06%. In 2021, it decreased to 0.03%, then slightly increased to 0.04% in 2022, and surged to 0.45% in 2023. The probability curve of jet fire accidents shows an initial decline followed by a rapid increase in recent years. The probability of a vapor cloud explosion accident occurring was 0.02% in 2020, 0.15% in 2021, 0.17% in 2022, and decreased again to 0.02% in 2023. The probability curve for vapor cloud explosion accidents initially decreased, then gradually increased. The probability of a toxic gas leakage accident occurring was 0.02% in 2020, 0.025% in 2021, 0.04% in 2022, and surged to 0.45% in 2023. The probability curve for toxic gas leakage shows a steady upward trend in recent years. According to Fig. 10b, the data is input into the Bayesian network for calculation. The probability of an accident risk level 3 was 8.1%, while the probability of level 4 was 11.4%. Based on the maximization principle, the accident level was determined to be 4, which aligned with the semi-quantitative evaluation method of the cloud model. The experimental data shows that the proposed method can express the risk evolution over different time periods, having a significant impact on preventing accidents and reducing the consequences based on accurate accident levels. To assess the effectiveness in large-scale data evaluation, comparative experiments were conducted using three evaluation methods: the scenario model evaluation method, the cloud-based accident evaluation method, and the Bayesian-based accident evaluation method, represented as F1, F2, and F3. The results of these experiments are shown in Fig. 11.

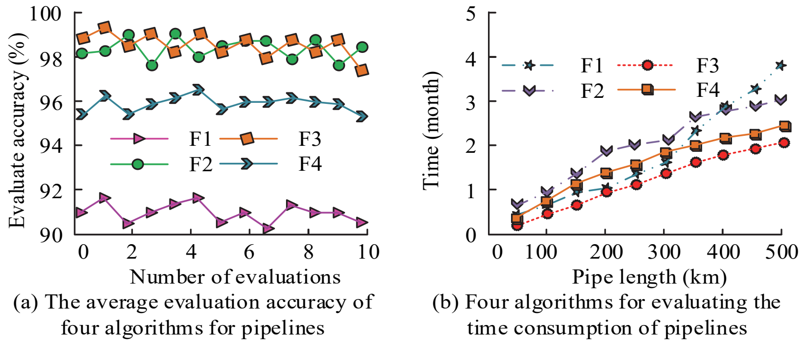

Figure 11 Pipeline evaluation of different models

According to Fig. 11a, the average evaluation accuracy for accident occurrence was 91.4% for F1, 98.2% for F2, and 98.8% for F3. The average assessment accuracy for F4 was 96.4%. In Fig. 11b, for the same pipeline, F1 took about 4 months to complete the evaluation, F2 took about 3 months, and F3 took about 2 months. F4 took about 2.5 months. The experimental data shows that the risk assessment method provides relatively accurate evaluation results. It is not only accurate for pipeline risk assessment but also demonstrates higher evaluation efficiency. The F3 method achieved the highest accuracy, showcasing its advantages in handling complex systems. The F4 machine learning algorithm showed higher accuracy compared to F1 and F2, but lower accuracy compared to F3. The F3 method also required the shortest time, reflecting its efficiency in information integration and processing. The F4 method utilized machine learning algorithms to capture complex risk patterns and potential abnormal behaviors through training on a large volume of historical data, thereby improving the accuracy of the risk assessment. However, the performance of machine learning models is highly dependent on the quality of the input data. Noise, missing, or mismatched historical data may lead to a decrease in the accuracy of the evaluation results.

A risk assessment method based on the cloud model and Bayesian network is proposed for safety management in an intelligent gas pipeline management system. According to consumer behavior theory, trust and transparency are crucial factors in evaluating corporate behavior. The research results demonstrated that the risk assessment method based on the Bayesian network and cloud model improved accuracy, reaching 98.8% and 98.2%, respectively, aligning with market expectations for efficient and transparent pipeline management. The results indicate that the Bayesian network-based method excels in both accuracy and efficiency. The average evaluation time for F3 was approximately 2 months. This enables enterprises to take appropriate preventive measures when potential risks arise, thus reducing the likelihood of accidents. However, the deviation from expectations was that, although the accuracy of the cloud model method reached 98.2%, its evaluation cycle remained relatively long, about 3 months. In some instances, this delay in decision-making could pose challenges. Adopting an efficient and reliable risk assessment model can significantly enhance the priority of pipeline management and maintenance. Based on the experimental data, the F3 Bayesian network model can effectively reduce the incidence of accidents, thereby decreasing threats to public safety. For stakeholders, encouraging pipeline operators to invest in advanced technologies such as Bayesian networks and cloud models is a practical measure to improve risk management capabilities, enhance response speed, and optimize accident prevention strategies.

5 CONCLUSION

This study introduces a Bayesian network-based risk assessment framework for intelligent gas pipeline management, addressing the limitations of traditional statistical and expert-driven evaluation methods. By integrating cloud-based uncertainty modeling and probabilistic reasoning, the proposed approach enhances the accuracy, efficiency, and adaptability of risk assessment. Experimental results confirm an accuracy rate of 98.8% and a two-month assessment period, demonstrating its superiority over conventional models.

Beyond its technical advantages, this method offers a scalable, data-driven approach for managing urban gas infrastructure, reducing risks associated with pipeline failures and enhancing proactive maintenance strategies. The findings highlight the significance of dynamic risk modeling, where Bayesian networks continuously update risk assessments by incorporating real-time data and historical trends.

However, some limitations remain. The approach relies on the availability and accuracy of historical data, and its computational demands may present challenges for real-time deployment on large-scale pipeline networks. Future research will focus on integrating IoT-enabled real-time monitoring and machine learning-based anomaly detection to enhance predictive capabilities. Additionally, hybrid AI-driven risk models combining deep learning with Bayesian inference could further optimize decision-making in intelligent infrastructure management.

By advancing predictive risk analytics, this research contributes to the development of safer and more resilient gas distribution networks, enhancing urban safety and operational efficiency in smart cities.

6 REFERENCE

[1] Ageed, Z. S., Zeebaree, S. R. M., Sadeeq, M. A. M., Abdulrazzaq, M. B., Salim, B. W., Salih, A. A., & Ahmed, A. M. (2021). A state-of-the-art survey for intelligent energy monitoring systems. *Asian Journal of Research in Computer Science*, *8*(1), 46-61. https://doi.org/10.9734/ajrcos/2021/v8i130192

[2] Yuvaraj, N., Praghash, K., Raja, R. A., & Karthikeyan, T. (2022). An investigation of garbage disposal electric vehicles (GDEVs) integrated with deep neural networking

(DNN) and intelligent transportation system (ITS) in smart city management system (SCMS). *Wireless Personal Communications*, *123*(2), 1733-1752. https://doi.org/10.1007/s11277-021-09210-8

[3] Eames, I., Austin, M., & Wojcik, A. (2022). Injection of gaseous hydrogen into a natural gas pipeline. *International Journal of Hydrogen Energy*, *47*(61), 25745-25754. https://doi.org/10.1016/j.ijhydene.2022.05.300

[4] Wu, X., Zhang, H., Yang, M., Jia, W., Qiu, Y., & Lan, L. (2022). From the perspective of new technology of blending hydrogen into natural gas pipelines transmission: Mechanism, experimental study, and suggestions for further work of hydrogen embrittlement in high-strength pipeline steels. *International Journal of Hydrogen Energy*, *47*(12), 8071-8090. https://doi.org/10.1016/j.ijhydene.2021.12.108

[5] Erdener, B. C., Sergi, B., Guerra, O. J., Chueca, A. L., Pambour, K., Brancucci, C., & Hodge, B. M. (2023). A review of technical and regulatory limits for hydrogen blending in natural gas pipelines. *International Journal of Hydrogen Energy*, *48*(14), 5595-5617. https://doi.org/10.1016/j.ijhydene.2022.10.254

[6] Alquier, P. (2024). User-friendly introduction to PAC-Bayes bounds. *Foundations and Trends® in Machine Learning*, *17*(2), 174-303. https://doi.org/10.1561/2200000100

[7] Luo, N., Yu, H., You, Z., Li, Y., Zhou, T., Jiao, Y., & Qiao, S. (2023). Fuzzy logic and neural network-based risk assessment model for import and export enterprises: A review. *Journal of Data Science and Intelligent Systems*, *1*(1), 2-11. https://doi.org/10.47852/bonviewJDSIS32021078

[8] Romano, M., Contu, G., Mola, F., & Conversano, C. (2024). Threshold-based naïve bayes classifier. *Advances in Data Analysis and Classification*, *18*(2), 325-361. https://doi.org/10.1007/s11634-023-00536-8

[9] Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*(1), 45-57. https://doi.org/10.1093/geronb/gby065

[10] Cerniauskas, S., Junco, A. J. C., Grube, T., Robinius, M., & Stolten, D. (2020). Options of natural gas pipeline reassignment for hydrogen: Cost assessment for a Germany case study. *International Journal of Hydrogen Energy*, *45*(21), 12095-12107. https://doi.org/10.1016/j.ijhydene.2020.02.121

[11] Liu, B., Liu, S., Guo, S., & Zhang, S. (2020). Economic study of a large-scale renewable hydrogen application utilizing surplus renewable energy and natural gas pipeline transportation in China. *International Journal of Hydrogen Energy*, *45*(3), 1385-1398. https://doi.org/10.1016/j.ijhydene.2019.11.056

[12] Yan, Y., Liu, D., Gao, B., Tian, G. Y., & Cai, Z. C. (2020). A deep learning-based ultrasonic pattern recognition method for inspecting girth weld cracking of gas pipeline. *IEEE Sensors Journal*, *20*(14), 7997-8006. https://doi.org/10.1109/JSEN.2020.2982680

[13] Juhong, W., Zhang, S., Xin, W., Ronguang, L., & Ting, W. (2020). Exploration and practice of data visualization for intelligent pipeline in China-Russia Eastern Gas Pipeline. *Oil & Gas Storage and Transportation*, *39*(2), 169-175.

[14] Qiu, Y., Zhou, S., Gu, W., Ding, S., Han, G., Zhang, K., & Lv, H. (2022) .Multi-stage flexible planning of regional electricity-HCNG-integrated energy system considering gas pipeline retrofit and expansion. *IET Renewable Power Generation*, *16*(15), 3339-3367. https://doi.org/10.1049/rpg2.12586

[15] Keysers, C., Gazzola, V., & Wagenmakers, E. J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788-799. https://doi.org/10.1038/s41593-020-0660-4

[16] Wickramasinghe, I. & Kalutarage, H. (2021). Naive Bayes: Applications, variations and vulnerabilities: A review of literature with code snippets for implementation. *Soft Computing*, *25*(3), 2277-2293. https://doi.org/10.1007/s00500-020-05297-6

[17] Jackins, V., Vimal, S., Kaliappan, M., & Lee, M. Y. (2021). AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *The Journal of Supercomputing*, *77*(5), 5198-5219. https://doi.org/10.1007/s11227-020-03481-x

[18] Itoo, F., Meenakshi, & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, *13*(4), 1503-1511. https://doi.org/10.1007/s41870-020-00430-y

[19] Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behaviour*, *5*(10), 1282-1291. https://doi.org/10.1038/s41562-021-01177-7

[20] Utami, E. (2022). Comparison of naive bayes classifier, k-nearest neighbor and support vector machine in the classification of individuals on twitter accounts. *Jurnal Teknik Informatika*, *3*(3), 673-680.

[21] Yu, S. Y., Malawade, A. V., Muthirayan, D., Khargonekar, P. P., & Al Faruque, M. A. (2021). Scene-graph augmented data-driven risk assessment of autonomous vehicle decisions. *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 7941-7951. https://doi.org/10.1109/TITS.2021.3074854

[22] Sadeghi, A., Jabbari Gharabagh, M., Rezaeian, M., Rezaeian, M., Alidoosti, A., & Eskandari, D. (2020). Fire and explosion risk assessment in a combined cycle power plant. *Iranian Journal of Chemistry and Chemical Engineering*, *39*(6), 303,311.

[23] Ke, D. A. N., Shuai, J., Guang, Y., Wei, A. N. G., Meng, E. N. G., & Zhang, H. (2020). Evaluation method for basic failure probability of oil and gas pipeline based on accident statistics in the United States and enlightenment. *Oil & Gas Storage and Transportation*, *39*(5), 530-535.

[24] Afzal, F., Yunfei, S., Nazir, M., & Bhatti, S. M. (2021). A review of artificial intelligence based risk assessment methods for capturing complexity-risk interdependencies: Cost overrun in construction projects. *International Journal of Managing Projects in Business*, *14*(2), 300-328. https://doi.org/10.1108/IJMPB-02-2019-0047

[25] Ayyildiz, E. & Taskin Gumus, A. (2021). Pythagorean fuzzy AHP based risk assessment methodology for hazardous material transportation: An application in Istanbul. *Environmental Science and Pollution Research*, *28*(27), 35798-35810. https://doi.org/10.1007/s11356-021-13223-y

[26] Davies, M., Jones, R. D. O., Grime, K., Jansson-Löfmark, R., Fretland, A. J., Winiwarter, S., & McGinnity, D. F. (2020). Improving the accuracy of predicted human pharmacokinetics: Lessons learned from the AstraZeneca drug pipeline over two decades. *Trends in Pharmacological Sciences*, *41*(6), 390-408. https://doi.org/10.1016/j.tips.2020.03.004

[27] Shi, X., Wong, Y. D., Chai, C., & Li, M. Z. F. (2020). An automated machine learning (AutoML) method of risk prediction for decision-making of autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *22*(11), 7145-7154. https://doi.org/10.1109/TITS.2020.3002419

[28] Zhou, X., Yu, J. & Huang, M. (2020). Evaluation method for effect of tunneling on underground jointed pipelines. *Chinese Journal of Geotechnical Engineering*, *42*(1), 181-187.

[29] Tatineni, S. & Chinamanagonda, S. (2021). Leveraging artificial intelligence for predictive analytics in DevOps: Enhancing continuous integration and continuous deployment pipelines for optimal performance. *Journal of*

Artificial Intelligence Research and Applications, *1*(1), 103-138.
[30] Kim, C., Chen, L., Wang, H., & Castaneda, H. (2021). Global and local parameters for characterizing and modeling external corrosion in underground coated steel pipelines: A review of critical factors. *Journal of Pipeline Science and Engineering*, *1*(1), 17-35. https://doi.org/10.1016/j.jpse.2021.01.010
[31] Mao, Y., Liu, S., & Gong, D. (2023). A hybrid technological innovation text mining, ensemble learning and risk scorecard approach for enterprise credit risk assessment. *Tehnicki Vjesnik*, *30*(6), 1692-1703. https://doi.org/10.17559/TV-20230316000447
[32] Li, M., Zhang, Z., Lu, M., Jia, X., Liu, R., Zhou, X., & Zhang, Y. (2023). Internet financial credit risk assessment with sliding window and attention mechanism LSTM model. *Tehnicki Vjesnik*, *30*(1), 1-7. https://doi.org/10.17559/TV-20221110173532

Contact information:

ChunXiao MEI
(Corresponding author)
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: chunxiaomei0725@126.com

JianXin TAN
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: jiahongyan_111@163.com

Hao LI
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: yynanj_123@163.com

JingTao CHANG
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: zhangxl198503@163.com

YiFan WANG
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: hujingyihb@163.com

YiWei LU
Hebei Gas Co., Ltd,
Shijiazhuang, 050011, China
E-mail: liuying_198911@163.com