Risk and reliability analysis for oil and gas pipelines using data-driven methods

by

Yinuo Chen

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering

University of Alberta

Abstract

Oil and gas pipelines (OGP) play a crucial role in sustaining the economy. With the increasing speed of pipeline network building, there is a corresponding growth in energy supply and demand. Nevertheless, numerous pipeline network safety operating issues arise during oil and gas transportation, such as corrosion failure leading to leakage, often resulting in fatalities, significant environmental damage, and economic losses. Therefore, it is necessary to conduct risk and reliability analysis to identify accident precursors and prevent accidents before they happen. It can also predict the pipeline degradation process, avoid major failure, and determine the priority of risk mitigation, optimize resource allocation. Various categories of OGP systems generate a large amount of data during operation. However, current studies have encountered challenges using these data to model risk and reliability effectively. Existing studies have constructed generalized models that ignore specific characteristics of different OGP systems. Furthermore, some of these studies failed to use appropriate data sources and inadequately addressed the uncertainty associated with input data, resulting in inaccurate results. In addition, the current models exhibit computational inefficiency and operational complexity.

Therefore, this thesis aims to utilize data sources from different OGP systems to develop more efficient and accurate data-driven models for risk and reliability assessment. This thesis fully considers the characteristics of different OGP systems and creates different risk and reliability analysis models in a targeted manner. The proposed models are more comprehensive. At the same time, the structure is simplified and the operation is more convenient, which can significantly improve the computational efficiency of the models while obtaining more accurate analysis results.

For pipelines where in-line inspection (ILI) can not be conducted, a novel method of cloud-variable weight function is proposed to analyze the pipeline's risk level and critical risk factors by establishing a pipeline risk assessment index system. The proposed method fully considers the uncertainty in the evaluation process, resolves the contradiction of existing methods to model the fuzzy concepts accurately, optimizes the

weight distribution, and obtains a more scientific and reasonable assessment result. For gas transmission system (GTS), a structure mapping method based on failure modes and effects analysis (FMEA) is proposed to form the GTS's object-oriented Bayesian network (OOBN) framework, making the model more user-friendly. An accident precursor identification approach is proposed based on the piecewise aggregate approximation-cumulative sum (PAA-CUSUM) algorithm, which can better discover the potential risks in real-time. The proposed method identifies process anomalies through monitoring data and analyzes the events and propagation patterns with the highest potential risk. For pipelines where ILI can be conducted, a finite element (FE) model is established. A reliability prediction method based on Residual Neural Networks (ResNet) that can directly map the magnetic flux leakage (MFL) inspection data to the pipeline's reliability is proposed. Pipeline defect effective area model, rather than those based on just depth, is effectively integrated with deep learning models. Moreover, an innovative approach for reconstructing the defect profile using a novel hybrid neural network to accurately and efficiently map three-axial MFL signals to the defects' 3-D profile is also proposed. It utilizes the neural ordinary differential equation (ODE) as a module within the neural network architecture, which can map the MFL signals to the spatial position of each point on the defective concave surface. Additionally, the proposed model incorporates the Fourier integration kernel to enhance computational efficiency.

The contributions of this study lay the foundation of for OGP potential risk discovery, risk control and rehabilitation, and pipeline digital integrity management. The proposed research can be extended to investigate the risk and reliability problems for OGP systems considering more complex events and situations in future work.

Preface

This thesis is an original work by Yinuo Chen. The research topics have been published or submitted for publication under the supervision of Dr. Zhigang Tian. The journal papers and conference papers are with Yinuo Chen as the lead author and Dr. Zhigang Tian as the corresponding author.

Chapter 3 has been published as Yinuo Chen and Zhigang Tian. “Risk assessment of buried gas pipelines based on improved cloud-variable weight theory”. Reliability Engineering & System Safety, 2022, 221: 108374.

Chapter 4 has been published as Yinuo Chen, Zhigang Tian, Rui He, and Yifei Wang. “Discovery of potential risks for the gas transmission station using monitoring data and the OOBN method”. Reliability Engineering & System Safety, 2023, 232: 109084.

Chapter 5 has been published as a conference paper: Yinuo Chen, Zhigang Tian, Haotian Wei, and Shaohua Dong. “Reliability assessment for corroded pipes based on MFL inspection”. ISSAT conference, August 2023, San Francisco, USA. Chapter 5 also has been published as: Yinuo Chen, Zhigang Tian, Haotian Wei, and Shaohua Dong. “Reliability analysis of corroded pipes using MFL signals and Residual Neural Networks”. Process Safety and Environmental Protection, available online 20 February 2024.

Chapter 6 brings up a submitted journal paper: Yinuo Chen, Zhigang Tian, Haotian Wei, and Shaohua Dong. “Reconstruction of 3-D pipeline defect profile based on MFL signals and hybrid neural networks”. The manuscript is submitted to Reliability Engineering & System Safety.

Acknowledgments

Words cannot express my gratitude to my supervisor, Dr. Zhigang Tian, for his invaluable patience and guidance, immense knowledge, and continuous support. And thank Dr. Tian for offering me an opportunity to pursue the Ph.D. degree, a significant milestone in my personal academic journey. I also could not have completed this journey without my defense committee, Dr. Ahmed Qureshi, Dr. Xinming Li, Dr. Weixing Chen, and Dr. Wujun Si, who generously provided helpful advice and insightful comments.

I am fortunate to benefit from the financial support provided by the CSC, U of A, and Mitacs, which enables me to access a highly conducive study environment despite my modest background. Many thanks to all my colleagues from the RRL Lab. Particularly, I'd like to thank Miss Yunqiu Zuo, Mr. Yifei Wang and Dr. Han Zhang. The times we spent studying, collaborating, and having fun together will never be forgotten. Thanks should also go to U of A. During the days of my study here, I have felt the warmth of humanistic care, equality, and respect.

I am also thankful to all my friends in Canada, China, and America, who have supported me through my Ph.D. time or even longer. Particularly, I'd like to thank Miss Yanran Wang for her constant companionship and unconditional support for any decision I make. Thank Dr. Wenkang Zhang for his passion for life and self-discipline, which always motivate and inspire me. Thank Miss Jiatong Ling and Mr. Haotian Wei for their kind academic and spiritual support.

Also, I would thank Dr. Shaohua Dong and Dr. Hang Zhang from China University of Petroleum – Beijing for their encouragement and helpful suggestions.

Special thanks to my favorite band, Mayday, for always giving me strength.

At last, I would like to extend my sincere thanks to my husband, Dr. Daoquan Zhu. Maintaining long-distance relationships can be challenging; however, his support, understanding, and endless love have been instrumental in helping me navigate through difficult circumstances. I also deeply appreciate my parents' eventual support, despite their initial lack of understanding. I know that this thesis stands as a testament to their unconditional love. Furthermore, I would like to thank my in-laws and my family for all the love they show me!

Table of Contents

List of Tables

List of Figures

List of Abbreviations

OGP	Oil and gas pipelines
ILI	In-line inspection
GTS	Gas transmission system
FMEA	Failure modes and effects analysis
OOBN	Object-oriented Bayesian network
PAA-CUSUM	Piecewise aggregate approximation-cumulative sum
FE	Finite element
ResNet	Residual neural networks
ODE	Ordinary differential equation
MFL	Magnetic flux leakage
EGIG	European Gas Pipeline Incident Data Group
PHMSA	Pipeline and Hazardous Materials Safety Administration
QRA	Quantitative risk analysis
HSE	Health, safety, and environment
FTA	Fault tree analysis
ML	Machine learning
DL	Deep learning
HAZOP	Hazard and operability analysis
ETA	Event tree analysis
BTA	Bow-tie analysis
PN	Petri net
AHP	Analytic hierarchy process
BN	Bayesian network
SCADA	Supervisory control and data acquisition
LSTM-AE	Long short-term memory and auto-encoder
OCSVM	One-class support vector machine
VMD	Variational mode decomposition
HBM	Hierarchical Bayesian Model

SVM	Support vector machine
SSA-CNN	Sparrow search algorithm and convolutional neural network
DOFS	Distributed optical fiber sensors
RF	Random forest
TPD	Third-party damage
BPNN	Back-propagation neural network
RBFNN	Radial basis function neural network
VT-CNN	Visual transformation-convolutional neural network
SwinYv5	Swin transformer backbone YOLOv5
CR-CNN	Cross-residual convolutional neural network
UT	Ultrasonic testing
EC	Eddy current
EMAT	Electromagnetic acoustic transducer
GFA	Gaussian feature approximation
ANN	Artificial neural networks
BNN	Bayesian belief network
TFN	Triangular fuzzy numbers
TZFN	Trapezoidal fuzzy numbers
MCS	Monte Carlo simulation
CPT	Conditional probability table
CM	Cloud model
VWT	Variable weight theory
FD	Fishbone diagram
ISMO	Improved spider monkey optimization algorithm
MTTF	Mean time to failure
LOPA	Layer of Protection Analysis
LNG	Leaky noisy-or gate
EM	Expectation maximum
PCA	Principal component analysis

FSU	Filtration and separation unit
PRU	Pressure regulating unit
FMU	Flow metering unit
RBD	Reliability block diagram
ALARP	As low as reasonably practicable
HCA	High-consequence areas
CNN	Convolutional neural network
LSF	Limit state function
RK4-Net	4^{th}-Runge–Kutta Net
CR	Cost rate
MSE	Mean square error
RT	Radiographic testing
3-D	Three-dimensional
ReLUs	Rectified linear units
IVP	Initial value problem
MSE	Mean absolute error

Chapter 1: Introduction

1.1 Research Background

1.1.1 Significance of pipeline risk and reliability analysis

Pipelines are called the economy's arteries since they are one of the most cost-effective and safe ways to transport oil, natural gas, and refined oil products [1]. They transport substantial quantities of oil and gas products from drilling sites to refineries, petrochemical facilities, and ultimately, to residential and commercial consumers, constituting an essential component of the energy industry. For example, the Canadian pipeline network comprises more than 840,000 *km* of transmission, gathering, and distribution pipes, including 117,000 *km* of large-diameter pipelines, with considerate pipeline infrastructure in most provinces [2]. In the United States, roughly 118,000 *km* of pipelines transport liquid petroleum, while over 1,491,000 *km* transport natural gas (including the distribution lines for households, offices, and businesses) [3].

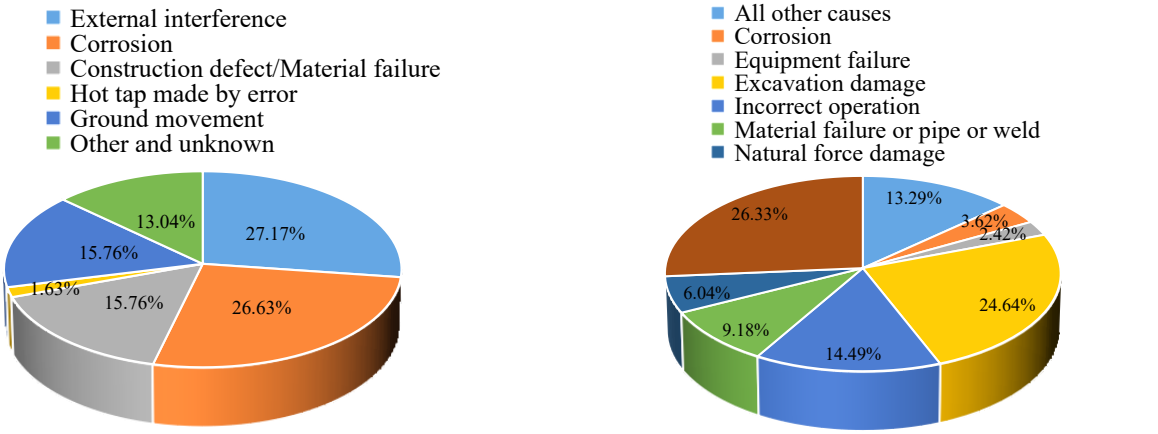
Due to the combustible and explosive nature of oil and gas media, pipeline accidents occur from time to time, resulting in dire environmental, societal, and economic consequences [4-6]. Biezma [7] summarized 23 OGP major accidents with 4,329 fatalities. According to statistics, there were 5598 major pipeline incidents in the United States between 1995 and 2014, with an average cost of 352 million USD. The overall cost of the accidents is estimated to be around USD 7 billion [8]. These statistics demonstrate that pipeline risk management is crucial to responsible development and a sustainable future.

Many reasons lead to pipeline failure and accidents [9]. Figure 1.1 (a) shows the European Gas Pipeline Incident Data Group (EGIG) database's five categories of pipeline failure causes and incident distribution per cause from 2010 to 2019 [10]. Figure 1.1 (b) shows the causes of serious gas pipeline incidents from 2005 to 2020

from the Pipeline and Hazardous Materials Safety Administration (PHMSA) [11]. PHMSA has established eight different categories of pipeline failure causes. The "excavation damage" category established by PHMSA is similar to the "external interference" category established by EGIG, both accounting for a large percentage of pipeline incidents. Other scholars have identified factors affecting pipeline safety when researching specific pipelines, as shown in Table 1.1. Thus, it can be found that risk of pipeline system include many categories, and in many cases, pipeline risk usually arises from the combination of multiple risk factors.

Table 1.1 Research studies on factors affecting pipeline safety

Research studies	Main factors affecting pipeline safety
[12]	Pipeline installation and backfill, equipment testing
[13]	Corrosion and external interference
[14]	Seabed soil; Man-made drilling oil stolen
[15]	Hot work with an open flame
[16]	Corrosion and external interference



(a) EGIG incident distribution per cause in the 2010–2019 period

(b) PHMSA incident distribution per cause in the 2005–2020 period

Figure 1.1 Incident distribution per cause of two databases

Therefore, it is necessary to conduct an overall risk assessment of the pipeline and specific analysis of key risk factors using multiple data sources. Risk and reliability assessment is the most critical part of risk management. It is low-cost and effective for pipeline operators to identify accident precursors and predict pipeline degradation processes to ensure safe pipeline operation. In complex systems, the results can provide

a basis for whether a component requires risk mitigation and identify the priority of risk mitigation. High-risk components may cause accidents, and control measures can be prioritized for their risks to prevent accidents before they happen.

Meanwhile, it is also essential to quantify the defects using inspection signals and assess the reliability of oil and gas pipeline systems. This enables more precise quantification of system performance, comprehension of reliability characteristics, execution of efficient maintenance strategies, reduction of maintenance expenses, and evaluation of inspection intervals.

1.1.2 Basic theories and definitions

The following are some basic definitions and concepts related to the research of this thesis.

(1) Risk

Decision-making is fundamentally based on the quantification of "risk." Risk analysis is a process that involves the use of qualitative and quantitative data to estimate and quantify the level of risk associated with a particular situation. Potential causes, consequences, and the likelihood of unexpected incidents should be considered during this process. The resulting information can then be used to support decision-making in scientific risk management. The NORSOK standard Z-013 [17] defines quantitative risk analysis (QRA) as four steps: risk estimation, risk analysis, risk assessment, and health, safety, and environment (HSE) management. ISO 31000:2009 [18] also defines four main steps in the risk management process: risk identification, risk analysis, risk evaluation, and risk treatment. Risk identification, also known as hazard identification, is an essential step before conducting risk analysis. Hazards can be classified into three categories [19]: (1) Time-dependent: external corrosion, internal corrosion, cracking, etc. (2) Time-independent (random): third-party damage, incorrect operation, and natural disasters. (3) Stable (resident): manufacturing and construction defects. There may be some variability in the expression of these criteria, but the definitions and the analysis framework are similar. Risk analysis was first used in nuclear safety studies in

the 1960s, and it was then expanded to other fields, such as the military and chemical industries [20]. Nowadays, risk analysis methodologies have developed rapidly in response to the growing demand for pipeline operation safety.

Due to the deterioration of the OGP and the continuous changes in operating conditions and environment, the inadequacies of traditional static risk analysis methods, such as fault tree analysis (FTA), failure mode and effect analysis (FMEA), etc., are exposed. Many researchers have worked on developing dynamic risk analysis methods in time dimensions to accomplish the dynamic description of the system's or a certain factor's risk changes [21, 22]. According to the literature search, most dynamic risk analysis approaches are based on upgrading traditional methods, with various data and information updating methods used to adjust the analysis results continually.

Because of the advances in sensor technology, OGP development is heading toward intelligence and digitalization. With the advancement of "Industry 4.0," "Oil and Gas 4.0" has just been added to the agenda [23]. During the operation of the OGP, a massive amount of data is generated. Hanga [24] stated that machine learning (ML) has enormous potential in oil and gas industry applications, especially regarding data analysis and interpretation. Mohammadpoor [25] also demonstrated how big data technology is becoming more prevalent in oil and gas transportation. As a result, many scholars established pipeline risk assessment and prediction models based on intelligent technologies such as machine learning (ML) and deep learning (DL) through collectible data to address safety issues. Therefore, risk analysis methods are no longer limited to traditional thinking, and their application in the OGP industry has entered a new era.

(2) Reliability

Both academic research and practical applications have achieved significant advancements in the structural reliability of pipelines. The term "structural reliability" refers to the likelihood that a given structure will provide adequate performance during its expected lifespan [26]. Reliability methods can be used not only to estimate the remaining useful life of a pipeline but also to guide pipeline operators in the repair and maintenance activities.

Research on pipeline reliability can be broadly divided into two categories: (1) assessment and prediction based on different data types. (2) assessment and prediction based on empirical models, which include deterministic models, such as linear corrosion rate model [27], and probabilistic models, such as gamma process-based corrosion growth model [28].

(3) Data-driven

A significant quantity of different types of data is produced in the routine operation of a pipeline. For pipeline operators, the concept of "data-driven" indicates identifying relevant analytical models based on factual information, data, and their features, then informing and guiding the decisions related to risk and reliability management. By acquiring quantifiable information via data analysis, data-driven approaches enable decision-makers to circumvent preconceived notions and partiality, leading to improved identification of potential pipeline risks and more precise reliability analysis outcomes.

Wen [29] summarized the current data-driven approaches into three categories: statistics-based, logic-based, and ML-based. The statistics-based is also the basis of other data-driven strategies. They are concerned with the data's surface features, such as statistical parameters or exploratory analysis of distribution. It is often used to conduct the initial analysis of OGP's risk and reliability [30]. Logic-based methods are typically founded upon a certain cause-effect logic or a total-division logic, and its configuration may be depicted in a logical chart, affording the advantages of clarity and comprehensibility. In the area of OGP, the most commonly used logic-based methods include FMEA [31, 32], hazard and operability analysis (HAZOP) [33], FTA [34], event tree analysis (ETA) [35], bow-tie analysis (BTA) [36-38], Petri net (PN) [39], analytic hierarchy process (AHP) [40], Bayesian network (BN) [41], etc. ML-based methods use emerging machine learning algorithms to discover potential information and patterns through deep data mining. Its specific advantages and disadvantages will be explained in detail in this thesis later.

(4) Uncertainty

Uncertainty is a concern that cannot be avoided in OGP's risk and reliability

assessments. In contrast to "certainty," the phrase "uncertainty" is used. Certainty refers to the analyst's ability to precisely and digitally characterize, specify, or forecast system behavior and other phenomena [42]. A lack of or incomplete knowledge often generates uncertainty, and it may be reduced to some extent as additional information becomes available. Many factors impact uncertainty, including the reliability of the quantitative data, the volume of data, and how closely the model assumptions match the real-world situation [43].

The process of achieving the description, characterization, and propagation of uncertainty, as well as the analysis of the outcomes, is referred to as uncertainty analysis. In risk analysis, there are two forms of uncertainty [44]: (1) variability and randomness induced by the system's stochastic behavior; (2) imprecision caused by measurement errors and a lack of information about the system. In OGP systems, the former, also known as aleatory uncertainty, arises primarily from the uncertainty inherent in the stochastic behavior of many physical processes, such as geological hazards, the unpredictability of third-party damage, and reduced reliability due to pipeline degradation. Aleatory uncertainty is unavoidable. For example, defects may produce different results when tested in the same setting several times, and this form of uncertainty can only be mitigated to a certain extent. It cannot be eliminated from the system [45]. The latter, often referred to as epistemic uncertainty, is primarily concerned with the amount of data and data quality used in risk and reliability evaluation in OGP systems. With the accumulation of information and the upgrading of data processing tools, this sort of uncertainty may be decreased.

1.2 Motivation

Some scholars have applied various data-driven methods to deal with the problems related to risk and reliability analysis. However, in the field of OGP, data-driven risk and reliability research needs to be further deepened and improved. Through the investigations, it is found that the following challenges still exist in the current studies. Some quantitative risk analysis studies are based on expert-inspired and fuzzy logic

theories. In recent research, related scholars [46, 47] have quantified the evaluation set given by decision-makers based on triangular fuzzy numbers or trapezoidal fuzzy numbers. Although the fuzzy operators can express the uncertainty of the weight calculation process, they appear in the final result as a specific value after defuzzification, still losing part of the uncertainty [48]. Moreover, the assignment of weights based on constant weight theory, such as AHP, is not scientifically rigorous and can easily lead to conclusions that are not consistent with reality.

On the other hand, due to the lack of data, the probabilities of the basic events required for traditional FTA, ETA, and BTA are obtained using expert inspiration, fuzzy logic, or by getting data directly from a common dataset [49, 50]. These data are not as close to reality as those obtained directly from operating conditions and can only be updated periodically for risk. Therefore, using these data not only reduces the accuracy of the assessment results but also does not allow for real-time risk analysis. In addition, FT and BT models have complex structures with numerous nodes and inconvenient operations, leading to inefficiencies.

In detecting defects in OGP and researching the reliability of pipelines with defects, the current approaches only focus on the depth of defects and ignore the defect's other geometry information, such as defect shape or profile, leading to less precise reliability analysis results. In addition, when using MFL signals for analysis, some of the existing studies only limit themselves to analyzing the single-axis MFL signals or use only the axial and radial signals as the basis for analysis. Thus, the model's accuracy is also not guaranteed. Also, most research focuses on processing MFL inspection data to extract the geometry of the defect and then utilizing the geometry information to conduct reliability assessment. The computation and analysis techniques are time-consuming and sophisticated. Consequently, pipeline operators may take significantly longer to acquire the final reliability result instead of effectively analyzing the pipeline state and executing maintenance steps immediately.

For different pipeline systems, such as the pipelines where ILI can be conducted and cannot be conducted, the gas transmission system, risk and reliability modeling can

be targeted based on the characteristics of these systems. Based on the research gaps illustrated above, when establishing the risk and reliability models, it is necessary to fill the gaps and obtain more accurate results.

1.3 Thesis Objectives

To address the above-mentioned challenges, this Ph.D. research focuses on developing effective data-driven models for risk and reliability assessment of different pipeline systems based on various data sources. Four sub-objectives are listed below.

(1) Develop a multi-factor coupled OGP risk assessment model and propose a novel weighting method for expert judgment, making the evaluation results more reasonable.

(2) Explore the potential risks of pipeline transmission system and develop a method to identify accident precursors in real-time through monitoring data.

(3) Develop a reliability prediction model based on deep learning methods that can directly map the MFL inspection data to the pipeline's reliability.

(4) Develop a deep learning model to directly reconstruct the 3-D profile based on the MFL inspection data.

1.4 Thesis Outline

The thesis is prepared following the guidelines from the Faculty of Graduate Studies and Research (FGSR) at the University of Alberta. The thesis, with seven chapters, is organized as follows.

Chapter 1 provides the research background of pipeline risk and reliability analysis. The research motivation is introduced. A brief statement on the thesis objectives and outline are also presented in this chapter.

Chapter 2 demonstrates a comprehensive analysis and categorization of academic publications with regard to risk and reliability assessment of oil and gas pipelines. This chapter summarizes the various data sources including monitoring data, ILI data,

historical records, expert judgment, and simulation data; outlines strategies to examine, analyze, and predict all major pipeline risks; analyzes the uncertainties inherent in the modeling process, and demonstrates the methods employed to address them.

Chapter 3 proposes a novel method based on cloud-variable weight theory to analyze the pipeline's risk level and critical risk factors by establishing a pipeline risk assessment index system. The proposed method fully considers the uncertainty in the evaluation process, resolves the contradiction of existing methods to model the fuzzy concepts accurately, optimizes the weight distribution, and obtains a more scientific and reasonable assessment result. The results of the case study illustrate that the proposed method is beneficial for helping pipeline operators determine the pipeline risk status and maintenance priority in the pipeline system.

Chapter 4 proposes a structure mapping method based on FMEA to form the OOBN framework for GTS. An accident precursor identification approach is also proposed based on the PAA-CUSUM algorithm. The proposed method identifies process anomalies through monitoring data and analyzes the events and propagation patterns with the highest potential risk. The case study results demonstrate that the proposed method is beneficial for assisting station operators in identifying possible hazards and providing a foundation for daily risk mitigation.

Chapter 5 proposes a reliability prediction method based on ResNet that can directly map the MFL inspection data to the pipeline's reliability. Pipeline defect effective area model, rather than those based on just depth, is effectively integrated with deep learning models. Case studies of FE simulations and industrial applications illustrate that the suggested approach is capable of assessing the reliability of corroded pipes in a more timely and accurate manner than traditional methods. The proposed method is also helpful for pipeline operators to understand the pipeline risk condition and obtain suggestions for optimizing costs and re-assessment intervals.

Chapter 6 proposes a novel hybrid neural network-based 3-D defect reconstruction method, which can directly inverse the defect shape. The three-axis MFL signals are used as model input. The neural ODE maps the MFL signals to the spatial position of

each point on the defective concave surface based on its excellence in parameterizing a homeomorphism between two sets. Furthermore, the model incorporates the Fourier integration kernel to enhance computational efficiency. Due to the difficulty in obtaining sufficient amounts of high-quality experimental data, the proposed model can be trained on data from FE simulations and then transferred to the experimental dataset, effectively solving the problem. The proposed method can guide pipeline operators to perform maintenance management and is helpful to establish the pipeline digital twin model.

Finally, Chapter 7 summarizes the research tasks completed in this thesis. The corresponding limitations and future research plans are also discussed in this chapter.

Chapter 2: Literature Review

2.1 Applicable methods for different data types

2.1.1 Pipeline base data

The basic pipeline information is shown in Table 2.1, which is a basis for analyzing other types of data. In the following sections, we choose the five most common data types for detailed analysis.

Table 2.1 Basic information

Category	Data
Attribute data	Wall thickness
	Seam type
	Diameter
	Manufacturer and manufacturing date
	Materials
	Buried depth
	laying method

2.1.2 Monitoring data

The primary focus of risk analysis utilizing monitoring data is typically centered on failure and consequence analysis, specifically concerning pipeline leakage. The widespread utilization of supervisory control and data acquisition (SCADA) systems facilitates the accessibility of pressure, flow rate, and temperature data during pipeline operation. Pipeline leakage is bound to result in anomalies in the monitoring data [51]. Therefore, this has garnered significant attention from scholars. Zuo et al. [52] transformed the pressure, flow rate, and temperature into the proper multivariate time series then used the trained long short-term memory and auto-encoder (LSTM-AE) to extract the data features and used the one-class support vector machine (OCSVM) to deal with the features and obtain the leakage threshold. Bhaskaran et al. [53] proposed

a method for detecting cracks and blockages through anomalous pressure fluctuations based on K-means clustering. Li et al. [54] extracted the actual signals from the pressure, flow rate, and temperature data of subsea gas pipelines by eliminating the noise based on variational mode decomposition (VMD) and used Hierarchical Bayesian Model (HBM) to obtain the failure of the process operations. Priyanka et al. [55] proposed a pipeline digital twin model in which abnormal pressure data is used to detect the failure precursor. Liu et al. [56] estimated the leakage orifice diameters of gas pipelines using BN and pressure and flow rate data. Aljameel et al. [57] compared five widely used ML algorithms and concluded that the support vector machine (SVM) is the best for analyzing the pressure, flow rate, and temperature data.

In addition to data from the SCADA system, some scholars proposed utilizing specific sensors or devices to monitor the pipeline risk and detect leakage. Lu et al. [58] and Yang et al. [59] used acoustic wave sensors to collect the signals and conducted the leak diagnosis based on SVM combined with other ML algorithms. Li et al. [60] utilized the sparrow search algorithm and convolutional neural network (SSA-CNN) method to analyze the negative pressure wave to detect the leakage. Wang et al. [61] collected the temperature and vibration data from distributed optical fiber sensors (DOFS) and identified the leakage based on the random forest (RF) model. Guerriero et al. [62] developed a dynamic BN model to fuse the data monitored from DOFS and improve the leakage alarm.

Due to the advantages of greater sensitivity and higher precision in location, optical fiber monitoring technology is increasingly used in a wide range of applications [63]. In the research area of risk identification, an optical fiber sensing system is always used to discover intrusion events or third-party damage (TPD) [64-66]. This system can provide both time-domain and spatial-domain data. Yan et al. [67] used a back-propagation neural network (BPNN) as the classifier to recognize human activities. Zhao et al. [68] introduced a novel approach for identifying spatiotemporal signals using artificial intelligence object detection, specifically YOLOv4, which can identify manual digging, mechanical excavation, and moving targets well. There are also other

types of monitoring data that may be used to identify the TPD events, such as the massive location data of mobile devices [69], the polarimetric synthetic aperture radar imagery [70], subsidence deformation data [71] etc., which need to be further researched.

2.1.3 ILI data

The monitoring of defect growth and assessment of pipeline structural integrity are commonly carried out through the application of ILI tools, which refers to intelligent pipeline integrity gauges (PIG), carrying various kinds of sensors to detect pipeline defects. Magnetic principle-based MFL tools are the most commonly utilized among the available ILI options [72]. This technology's challenge is measuring the MFL field, distinguishing defects from non-defective signals like welds and tees, accurately quantifying the defects from the collected magnetic flux signals, and analyzing the pipeline's reliability using the quantified and reconstructed defect profiles. Over the last two decades, significant technological advancements have enabled scholars to effectively address the challenges associated with improving the MFL signals' quality and recognizing defects [73-80]. Recent studies have focused on the estimation of defect sizes, defect reconstruction, as well as reliability assessment through defect parameters obtained from MFL.

In the defect quantification and reconstruction research area, Joshi et al. [81] used a data set containing 71 samples with different geometries and radial basis function neural network (RBFNN) to obtain the defects' 3-D depth profile. Chen et al. [82] also used RBFNN to reconstruct the two different shapes of the defect 3-D profile. Ma et al. [83] improved RBFNN by the immune algorithm, which can size the pipeline corrosion. Kandroodi et al. [84] constructed a database including 337 simulated defects and 31 experimental samples and proposed a 2D image processing method to obtain the defect width. Mohamed et al. [85] studied different neural network structures for defect length and depth estimation, used the Levenberg-Marquardt back-propagation learning algorithm to train the data, and found that the neural networks exhibited optimal

performance, achieving the accuracy of 86% and 89% for error tolerances of ±10% and ±15%. Lu et al. [80] established a magnetic dipole model to obtain the data and used a visual transformation-convolutional neural network (VT-CNN) to estimate the defect size. Layouni et al. [86] proposed an artificial neural networks-based method to recognize the defect length and depth. Wang et al. [87] collected 479 defect samples and integrated prior knowledge into the neural networks to quantify the defect sizes. Yuksel et al. [88] used the MFL signal interpolation method to increase the sample size from 100 to 1000, then combined the Swin Transformer Backbone YOLOv5 (SwinYv5) algorithm with the cross-residual convolutional neural network (CR-CNN) for defect detection and quantification, which is claimed to have a precision of 98.9%.

In addition to MFL, advanced ILI technologies include ultrasonic testing (UT), eddy current (EC), and electromagnetic acoustic transducer (EMAT) [89-91]. By using these techniques, there are also some data-driven methods to estimate the defect size. For example, Pyle et al. [92] introduced a Gaussian feature approximation (GFA) method to improve the explainability of the neural networks and effectively quantify the crack-like defects based on UT. Xiong et al. [93] proposed an improved deep extreme learning machine method using eddy current test data to predict the defect size accurately. Yan et al. [94] combined CNN with SVM to recognize the girth weld cracking using EMAT detection data.

The reliability analysis using data-driven methods may be conducted based on the inspected defect information. Zhang et al. [95] modified the safety factor parameter based on the pipeline data mining methods, enabling the SF obtained from various factors instead of only the pressure, providing a new reliability analysis perspective. Xiang and Zhou [96] developed a dynamic BN to model the corrosion growth and finally obtained the failure probability. Zhang and Tian [97] used artificial neural networks (ANN) to perform the reliability assessment of pipelines with multiple corrosion defects. Anghel [98] proposed a reliability classification procedure using a minimax SVM algorithm to obtain more accurate results. Adumene et al. [99] considered the interconnections of corrosion factors and failure modes and proposed a

novel reliability prediction method based on the defect parameters. Shabarchin et al. [100] proposed a Bayesian belief network (BBN) that integrated various corrosion and failure pressure models to calculate the failure probability. Woldesellasse and Tesfamariam [101] combined the BBN with GIS technology, enabling pipeline operators to estimate the failure probabilities of pipelines located in specific areas.

The resolution of a pipeline MFL detector refers to the smallest size of magnetic defects it can detect. For instance, if an MFL detector has a resolution of 1 millimeter, it means it can detect magnetic defects with a diameter of 1 millimeter. Higher resolution allows the detector to identify smaller and more subtle magnetic anomalies, which is crucial for identifying potential issues within the pipeline. Generally, the resolution of a pipeline MFL detection system directly impacts its performance in early detection of pipeline issues and ensuring pipeline safety. Low resolution in a pipeline MFL detector may introduce uncertainties in defect quantification. For example, low resolution may prevent the detector from accurately measuring and estimating the size of smaller defects within the pipeline, leading to uncertainties in quantifying defect sizes. In situations with low resolution, the detector may be more prone to false positives, incorrectly labeling some non-defective sections of the pipeline as defective. Additionally, inadequate resolution may make it challenging for the detector to capture and understand the complex shapes of defects, affecting the accurate quantification of defect shapes. The material and hoop stress of pipelines may also impact the performance of MFL detectors, particularly in terms of resolution and measurement accuracy. For instance, different pipeline materials exhibit varying responses to magnetic fields, potentially influencing the detector's ability to identify and locate defects. Additionally, the wall thickness and stress of pipelines can affect the propagation of MFL signals and the response of MFL tools. Larger wall thickness may weaken the magnetic flux leakage signal, while high stress may lead to signal distortion [102]. These factors may also introduce some level of uncertainty. As this thesis predominantly relies on simulated signals rather than real MFL tool data, the uncertainties arising from practice scenarios have not been extensively studied.

Investigating uncertainties caused by real MFL data will be a valuable research direction in the future, especially with the acquisition of a large volume of real data.

2.1.4 Historical records

Historical records typically include all recorded information related to pipeline operation, such as historical failure data, experimental data, operational records, and data sourced from published literature and well-known databases. For example, Guo et al. [103] developed a BN to identify the risk of TPD based on recorded accidents. Wen et al. [104] used the historical landslide data and other basic environmental data as input, and established a hybrid ML method to propose a risk assessment model for landslides. Yin et al. [5] and Shan et al. [105] introduced correction coefficients to modify the historical failure probability to obtain more accurate QRA results. Mazumder et al. [106] utilized a dataset consisting of 92 instances of documented pipe bursts to establish ML-based models. The results indicated that XGBoost is the most effective algorithm for predicting pipeline failure. Yang et al. [107] collated environmental data from the GIS system, which was used as the basis for risk analysis along with the pipeline's basic data, and finally established a Graph Convolutional Network (GCN) and clustering algorithm-based accident consequence assessment model. Kumari et al. [108] derived nonhomogeneous failure rates from the PHMSA database and used BN and ANN to conduct a risk analysis. Alves et al. [109] collected Brazilian historical failure data and established a bow-tie model to deal with pipeline theft. Wang and Li [110] proposed an unsupervised risk analysis algorithm to evaluate risk levels by utilizing a limited amount of historical failure records. Ma et al. [111] collected 314 data pieces from the literature and proposed a hybrid method to predict burst pressure.

Sometimes, due to the limited sources or amount of historical records, when establishing the risk analysis model, scholars often used historical data along with expert judgments, simulation data, etc., as input to the model [41, 112]. Examples include using historical records and expert judgment as BN's input [21, 113, 114] and ANN's input [115], using historical experimental data and simulated data to develop a

deep neural network (DNN) to predict failure pressure [116].

2.1.5 Expert judgment

For OGP operational risk analysis, expert judgment is necessary. For example, Alves and Lima [117] invited 180 pipeline experts with different backgrounds and established a Brazilian pipeline database based on expert elicitation. In addition, some particular situations may arise, such as the presence of a non-piggable pipeline that is buried underground. In these circumstances, the pipeline operators may be reluctant to incur significant costs associated with excavation to assess the pipeline's risk status. Consequently, it becomes essential to engage experienced experts well-versed in the pipeline's construction history, daily operation, management, etc., to provide their evaluations. Sometimes experts are also needed to give their opinion on consequential events, damage to assets, etc. An efficient way to accomplish the evaluation task is by developing a linguistic assessment criterion, such as utilizing a series of rubrics comprising labels: "very high," "high," "moderate," "low," or "very low," or alternatively, a basic numerical value or interval [118].

Many scholars have utilized expert judgments as input data for the models of OGP's risk assessment, as summarized in Table 2.2. Within the categories of data-driven models that are based on expert opinions, a minority of models use machine learning techniques, while the majority rely on logic-based approaches. The FTA and AHP-based indicator system models are prevalent among scholars. These two methods yield similar outcomes, and researchers are continually enhancing and refining them with technological advancements. The Bow-tie model is a hybrid of FTA and ETA, which can be used to start the analysis from the cause of the accident to the consequence more clearly and effectively by using logic diagrams. BN was introduced with the growing scholarly interest in pipeline dynamic risk assessment. One possible approach is to establish a direct correspondence between the configuration of a BN and FTA and Bow-tie models. The utilization of BN enables the improvement of the original static model to a dynamic one, thereby facilitating a more comprehensive depiction of the dynamic

changes of risk or the accident evolution process across different periods.

Table 2.2 Literature summary

Literature	Risk analysis model	Expert judgment expressions	Results achieved
[119]	Bow-tie	Triangular fuzzy numbers (TFN) of eleven and five linguistic terms	Obtained the failure probability of the top event and identified the critical factors
[46]	FTA	Trapezoidal fuzzy numbers (TZFN) of five linguistic terms	
[120]	FTA	Same as above	
[121]	FTA	TZFN of five and seven linguistic terms	
[21]	Bow-tie, BN	TFN of seven linguistic terms	
[122]	Petri net	Intuitionistic fuzzy evidential reasoning method	
[123]	Risk matrix, bow-tie	TZFN of five linguistic terms	
[124]	Bow-tie, DBN	TZFN of nine linguistic terms	
[125]	FTA, BN	Same as above	
[126]	Dynamic object-oriented Bayesian network (DOOBN)	TZFN of seven linguistic terms	Derived a comprehensive evolution path of pipeline accident and obtained the weight ranking of the consequences
[127]	Event tree analysis (ETA), decision-making trial and evaluation laboratory (DEMATEL), BN, Interpretative Structural Modeling (ISM)	Direct influence matrix	
[128]	DBN	TZFN of seven linguistic terms	
[129]	ML algorithms including multilayer perceptron (MLP), support vector regression (SVR) and random forest (RF)	Fuzzy logic inference model	Recognized the critical indexes and concluded that the RF model had the best performance
[130]	Risk assessment index system based on analytic hierarchy process (AHP)	Cloud model	Identified key risk factors, determined the risk level and risk priority, recognized the weaknesses of the pipeline management, and provided risk management suggestions
[131]	Vlsekriterijumska optimizacija i kompromisno resenje (VIKOR), criteria importance through inter-criteria correlation (CRITIC)	Decision matrix	
[132]	Failure mode and effects analysis (FMEA), VIKOR	Cloud model	
[133]	Improved FMEA, Grey Relations Theory (GRT)	TFN of five linguistic terms	
[134]	AHP	Cloud model	
[135]	AHP, TOPSIS	TFN of five linguistic terms	
[39]	Petri net	Cloud model	

With regard to the expression of expert opinions, a majority of the literature uses linguistic term sets based on TFN or TZFN to convey the inherent ambiguity of expert judgments. These term sets may consist of 5, 7, 9, or 11 terms. Furthermore, cloud modeling has gained significant popularity over the past ten years. The analysis of uncertainty in the expression of expert knowledge is specified in Section 2.2.

It should also be noted that the challenge of dealing with unreliable or subjectively biased expert views is common when we use expert judgments for risk analysis. To address this issue, the following strategies can be considered: (1) diverse expert selection. By choosing experts from different backgrounds, positions, and experiences, such as professors from universities, frontline workers, and managers from pipeline companies, diverse professional perspectives can be obtained. Different types of experts can provide opinions from unique angles, making the risk assessment more comprehensive. (2) Involvement of multiple experts. Increasing the number of experts helps reduce the impact of individual expert subjectivity on the overall evaluation. The convergence of collective wisdom may be more accurate and objective, as differences between various experts can offset each other, enhancing the robustness of the overall assessment. (3) Expert training. Providing expert training ensures that they fully understand the goals, methods, and standards of risk analysis. Training helps experts better comprehend their respective fields, enhancing their professional competence and judgment accuracy. It also emphasizes the importance of consistency, promoting uniformity among experts during the evaluation process. (4) Utilizing technical means such as the Delphi method. The Delphi method involves multiple rounds of anonymous expert surveys to achieve consensus. Through iterative cycles of anonymous expert discussions and voting, the Delphi method gradually reduces differences of opinion, aiming to reach expert consensus.

The comprehensive application of these strategies helps minimize uncertainty and subjective bias in expert opinions, improving the credibility and accuracy of risk analysis. By establishing a more diverse and multidisciplinary team of experts, combined with technical means like the Delphi method, it becomes possible to better

address challenges arising from individual expert differences and incomplete information, providing a more reliable foundation for risk assessment. In topic 1 of this thesis, three experts from different backgrounds (a professor with high academic attainments, an experienced engineer, and a senior manager of the pipeline company) provide their comments. They are well pre-trained, and the Delphi method is employed in the evaluation process.

2.1.6 Simulation data

In numerous instances, field data may not be accessible due to factors such as commercial confidentiality, geographical limitations, and exorbitant expenses. Consequently, simulated data obtained through diverse methods can substitute for actual data, thereby establishing a foundation for the risk and reliability management of pipelines.

The FE method is an effective tool for simulating data and is often combined with the ML model to obtain the results. Chen et al. [136] and Zhang et al. [137] used the FE method to establish a model of pipeline corrosion defect and obtain the burst pressure data, which is used to train the neural network. Xu et al. [138] simulated the pipeline failure caused by third-party excavation and then used ML algorithms to predict failure conditions. Peng et al. [139] established a FE model to analyze the pipeline's seismic response and finally obtain reliability based on FTA. Jiang and Dong [140] developed a nonlinear numerical FE model to simulate the falling objects impacting and conducted the probabilistic risk analyses.

Furthermore, as a simple but helpful tool, Monte Carlo simulation (MCS) is also very popular among scholars. Ossai et al. [141] and Timashev et al. [142] estimated the pipeline corrosion growth and reliability using MCS methods. Park et al. [143] used MCS to calculate the accident probability. As a basic simulation method, some methods are compared to MCS and are proven to be effective. The most typical one is subset simulation [144-146]. With guaranteed accuracy, the subset simulation is demonstrated to be far more efficient than MCS.

2.2 Uncertainty Processing Methods

Uncertainty is an intrinsic characteristic of risk, and quantifying the uncertainty of unexpected events is known as quantitative risk analysis [43]. This study concentrates on how to deal with uncertainty in data derived from various sources in light of the epistemic and aleatory uncertainties expounded upon in subsection 3.1.4. Currently, for dealing with the uncertainty related to pipeline risk and reliability analysis data, the common theories include probability analysis, interval analysis, fuzzy set theory, cloud models, Bayesian networks, etc.

(1) Probability analysis

Probability is a highly effective mathematical technique to describe and define uncertainty. This is due to the fact that probability serves as a statistical representation of the inherent unpredictability associated with random variation. Frequentists regard probability as a characteristic of the physical world, whereas Bayesian statisticians view probability as a personal conviction of the assessor, leading to the term “subjective probability.” The utilization of frequentist probabilities is the favored approach in characterizing the uncertainty of stochastic processes. However, in cases where stochastic processes involve cognitive uncertainty, the derivation of frequency probabilities is unattainable. As a result, subjective probabilities become a viable alternative [42]. The boundary between epistemic uncertainty and aleatory uncertainty is not well-defined, so sometimes decision-makers combine both types of uncertainty in their assessments due to the inability to discern the proportion of uncertainty attributable to each [45].

The probabilistic methodology for characterizing uncertainty is to consider a parameter as a stochastic variable, and its probability distribution can be represented by various probability density functions [147]. For example, Li et al. [27] used normal distributions to describe the uncertainty of parameters related to pipe corrosion failure. Fenyvesi et al. [148] and Salama et al. [149] employed normal distribution to measure the uncertainty caused by MFL and UT tool errors, respectively. Singh and Markeset

[150] combined the probabilistic approach and fuzzy set theory to describe the uncertainties of defect parameters obtained from ILI. Wu et al. [151] also combined the probability theory and possibility theory to process the parameter uncertainties during the reliability analysis. Dann and Maes [152] analyzed the inspection measurement error and derived the actual metal loss per ILI based on several probability distributions. We can also find from the literature that parameter uncertainties characterized by probability density functions can be propagated in the analytical model using Monte Carlo simulation to obtain the probability density distribution of the output results. The Monte Carlo simulation is popular due to its adaptability, straightforward implementation, and capacity to operate independently of model intricacy, multidimensionality, and nonlinearity. Additionally, it can take advantage of correlation information between variables [97, 136].

(2) Interval analysis

In cases where an evaluator possesses complete knowledge of a parameter's upper and lower bounds but has limited additional information, the uncertainty can be represented by intervals [153]. When interval probabilities are used to characterize uncertainty, interval arithmetic can be used to propagate the uncertainty of the input parameters in the analytical model to obtain possible bounds on the output results. For instance, Yu et al. [154] utilized interval analysis to quantify and propagate uncertainty in the pipeline risk analysis model based on AHP. Interval analysis is not only distinguished by its compatibility with human cognitive ability but also by its validity in handling uncertainty from any source and of any nature. Nevertheless, as the level of complexity in interval arithmetic escalates, the intervals tend to expand, and exhibit decreased precision in yielding the outcomes [155]. Furthermore, the outcomes are still intervals whose distributions exhibit identical levels of confidence and can not provide sufficient guidance for making decisions regarding the ultimate management of pipeline risk assessment.

(3) Fuzzy set theory

In many cases, expert knowledge is frequently employed as a substitute for

objective data in assessing risks, especially due to the challenges associated with acquiring precise probability functions. Fuzzy sets serve as a viable approach for representing uncertainty in natural language, as evidenced by the literature presented in Table 3. The fuzzy mathematical theory proposed by Zadeh [46] offers a potential solution to the limitations inherent in traditional probability theory, particularly in its ability to represent uncertainty effectively. Fuzzy set theory is a mathematical tool that characterizes linguistic terms as fuzzy numbers and employs continuous membership functions, such as triangular or trapezoidal fuzzy numbers, to represent uncertainty quantitatively. The utilization of fuzzy set theory in addressing uncertainty is considered a preferable approach compared to both probability analysis and interval analysis. The integration of fuzzy mathematical theory and risk analysis effectively mitigates the impact of uncertainty on the outcomes of risk and reliability assessments to a certain degree.

The relative scholars developed the fuzzy FT model, fuzzy bow-tie model, etc. For example, Shahriar et al. [119] developed a bow-tie model and utilized fuzzy logic to obtain probability data of basic events in FT and finally obtain the probability of the top event. Singh et al. [156] developed a fuzzy FT model and used the weakest t-norm calculations to reduce uncertainty accumulation. In addition, the literature presented in Table 3 demonstrates the widespread acceptance and utilization of fuzzy set theory. It enables analysts to derive informative conclusions about risk prevention and control, even when specific objective data is lacking.

(4) Cloud model

The cloud model theory is founded upon the principles of fuzzy mathematics and stochastic mathematical algorithms. It enables the integration of qualitative analysis and quantitative calculation to establish a corresponding mapping relationship and is an emerging uncertainty research method [157]. Wang et al. [158] used cloud model-based expert judgment to represent the reliability of the multi-state pipeline system. Liang et al. [159] integrated interval scoring and normal cloud model to handle the uncertainty in the risk assessment of long-distance pipelines. Similar to the application of fuzzy

numbers, cloud theory in the application of risk analysis is also first formed into a set of linguistic assessment criteria consisting of several cloud models, followed by experts carrying out the evaluation of ranks and weights. In contrast to fuzzy set theory, clouds consist of cloud drops that create a one-to-many mapping image rather than a distinct affiliation function fold line. Clouds have the ability to more effectively convey and disseminate uncertainty by utilizing their numerical features, including expectation, entropy, and hypertrophy, as well as employing more comprehensive operational guidelines.

(5) Bayesian theory

Bayesian theory is a probabilistic analysis technique commonly employed in risk and reliability analysis to effectively handle situations involving uncertain information [160, 161]. Bayesian networks utilize the chain rule to model the causal connections between random variables and account for the conditional dependencies among these variables. Each node represents each type of variable, the conditional probability table (CPT) assigned to each node represents the conditional dependencies, and the arcs represent the direct causal relationships between the linked nodes. As the engineering system progresses through its life cycle, the initial prior distribution of the system's nodes can be revised based on accumulating information. This updating mechanism enables deriving a posterior probability distribution that more accurately reflects the current state of events. Xiang and Zhou [96] established a dynamic BN which integrated corroded pipeline reliability analysis and uncertainty of ILI tools. Li et al. [162] developed a copula-Bayesian method to describe the complex connections of the risk variables. Qin et al. [163] considered measurement error and imperfect detectability of ILI tools using Bayesian Bayesian updating. Pesinis and Tee [164] constructed the corrosion growth model based on the hierarchical Bayesian method, which considered measurement errors. Heidary and Groth [165] established a hierarchical Bayesian model based on a non-homogeneous gamma process and processed the uncertainty of ILI data. Numerous studies have shown that Bayesian networks have also become a mainstream method for uncertainty handling in the field of risk and reliability

assessment.

2.3 Summary

This chapter mainly summarizes the applicable methods for different data sources and uncertainty processing methods. The motivations of Chapters 3 to 6 are based on this chapter.

For pipelines where ILI cannot be conducted, SCADA systems primarily monitor the overall operation of the entire pipeline system but may not provide adequate monitoring of localized pipe section anomalies. Therefore, expert judgment is often used as the main data source. The primary challenges are the quantification of the expert judgment, the language uncertainties, and the weight assignment method of experts' comments. These challenges are addressed in Chapter 3. The cloud-variable weight theory is proposed to handle the uncertainty and optimize the weights.

For GTS, the SCADA system can comprehensively collect data such as pressure, flow, etc. It monitors the operation of each critical component in GTS and stores the data within a specific time frame. These data serve as a solid foundation for analyzing potential risks. Chapter 4 primarily utilizes data from the SCADA system as the primary data source. Furthermore, historical data is used to acquire model parameters. The proposed method offers a feasible solution to the issue of current research relying on expert review and public datasets, which are incapable of conducting real-time risk analysis and generating accurate results. For uncertainty analysis, the proposed method uses BN and leaky noisy-or gate (LNG) to measure the uncertainty associated with risk factors.

For pipelines where ILI can be conducted, analyzing pipeline reliability based on ILI data is very straightforward and effective, and MFL is the most widely used ILI method. Due to business confidentiality and data volume limitations, Chapters 5 and 6 mainly use the simulated MFL data with a small amount of real data for validation. Based on Section 2.1.6, it can be found that the simulated data are also widely used in the field of pipeline risk and reliability and can be a good alternative to real data for

relevant analysis. Based on the information in Section 2.1.3, the existing methods can not directly get pipeline reliability from MFL signals, and most of the current methods focus on the depth quantification of defects and ignore the 3-D profile of defects. Chapter 5 proposes a deep learning method to map the MFL signal directly to the pipeline reliability trend. Chapter 6 solves the problem of reconstructing the 3-D profile of regular defects under arbitrary rotation angles.

In addition, it is necessary to note that, for different types of pipelines, we need to consider their distinct characteristics and operating environments to identify the most suitable risk analysis methods. For instance, when it comes to gas pipelines, urban buried gas pipelines and long-distance, large-diameter, high-pressure gas pipelines are common types. A risk analysis method based on expert judgment may be more appropriate for urban buried gas pipelines. These pipelines typically have smaller internal diameters, making conventional internal inspections challenging. Moreover, additional constraints, such as the complex urban environment around the pipelines, may make data collection relatively difficult. Therefore, relying on the knowledge and experience of experts may be a more feasible approach. In contrast, long-distance, large-diameter, high-pressure gas pipelines, characterized by larger diameters and higher pressures, are better suited for risk and reliability analysis methods based on internal inspection data. The operating environment and technical requirements of such pipelines make advanced techniques like internal inspections more feasible. These data can provide more detailed information about the pipeline's condition, supporting more precise assessments. Hence, for different pipeline types, it is essential to use tailored data and employ specific risk and reliability analysis methods accordingly.

Chapter 3: Risk assessment of buried gas pipelines based on improved cloud-variable weight theory

3.1 Introduction

The severity of the consequences of pipeline accidents and the diversity of causes of pipeline failures underscores how important it is to perform risk management of pipelines. However, city gas pipelines are located throughout the urban centers. In many cases, it is difficult to effectively perform combustible gas monitoring due to the pipeline geography and engineering practice limitations. At the same time, most city gas pipelines have diameters too small to allow for internal inspection. Therefore, the safety of pipelines cannot be effectively guaranteed if relied on these approaches alone. Based on various risk factors, an overall assessment of the risk status of the pipeline is necessary. Risk assessment is the most critical part of risk management. It is low-cost and effective for pipeline operators to ensure safe pipeline operation [116]. In complex systems, risk assessment results can provide a basis for whether a component requires risk mitigation and identify the priority of risk mitigation. High-risk components may cause accidents, and control measures can be prioritized for their risks to prevent accidents before they happen [166].

Many studies have been conducted about pipeline risk assessment. Examples include risk studies based on bow-tie models [167, 168], risk assessment cluster models based on data-driven methods [110], integrated failure probability estimation based on structural integrity analysis and failure data [169], and a comprehensive approach to risk management including considerations of frequency, gas release rate, welding defect analysis, et al. [170]. However, these methods often require precise probability values to support the evaluation, and in practical applications, there are often many limitations.

As stated by Vairo [171], due to incomplete statistics and knowledge, there will be some uncertainties on the likelihood and interdependence of root risk events in Fault Tree (FT) and events in Event Tree (ET), which lead to unrealistic results. According to Alves et al. [117], the lack of representative data is a major challenge for pipeline companies. In the absence of data, it is not possible to adequately estimate the probability and consequences of possible pipeline failures, thus adding much uncertainty to the risk evaluation of pipeline operations.

Therefore, to better solve practical problems in case of missing data, many researchers used methods based on expert decision-making to obtain more credible risk assessment results. However, due to human thinking patterns and cognitive styles, there are still uncertainties surrounding even experts making decisions. In addition, when multiple factors influence the decision outcome, the way the weight of each factor (i.e., the degree to which a factor contributes to the attribute value of its parent factor) is set is an issue worth studying. Table 3.1 is a summary of the relevant research in this area.

Table 3.1 Some relevant studies

Literature	Risk analysis method	Uncertainty processing methods	Factor weight processing method
[159]	FT	Cloud model (CM)	Structure entropy weight method
[134]	Risk assessment index system	Cloud model (CM)	Hierarchical analysis method (AHP)
[46]	FT	Fuzzy set theory	-
[47]	Risk assessment index system	Fuzzy set theory	-
[172]	Risk assessment index system	-	Decision-Making Trial and Evaluation Laboratory
[154]	Risk assessment index system	Interval analysis	AHP
[173]	Vulnerability indexes system	-	Machine learning methods

The following can be found from the research presented above: (1) The index system and FT are the most widely used multidimensional risk evaluation methods, collecting, assessing, and analyzing different factors to generate comprehensive results. Moreover, both CM and fuzzy set theory are good ways to handle uncertainties. (2) In some literature, risk index systems are established, and some constant weight theories such as AHP are used when calculating each index's weight. In this process, the weights of the indexes remain unchanged regardless of the changes in the attribute values of the

evaluated indexes. Therefore, when the attribute value of an index is far different from the normal value, but its weight value is very small, its impact on the overall risk assessment result will usually be ignored, leading to a less reliable result. (3) When expressing the uncertainty in the evaluation process, fuzzy operators such as trapezoidal fuzzy numbers are used to convert fuzzy concepts into exact values or intervals, thus losing part of the uncertainty in the conversion. It may still lead to unreliable assessment results. Therefore, fuzzy set theory is not as precise as the CM in expressing uncertainty. However, in some literature, the CM only expresses the results and is not reflected in the evaluation process. The authors use exact numbers throughout the evaluation process but only replace the numbers with a CM when expressing the results, which does not achieve the desired expression of uncertainty.

In this chapter, a novel cloud-variable weight theory is proposed to conduct pipeline risk assessment research. Compared with the constant weight theory mentioned above, the variable weight theory (VWT) can highlight the negative effects of the index with higher risk, which can obtain more reasonable results. CM is used to improve the calculation in VWT, and it is applied throughout the entire risk evaluation process. The uncertainty can also be better conveyed through the calculation between CMs, making the expression of uncertainty more adequate and accurate. Finally, a pipeline located in China is taken as an example. The pipe sections with higher risk levels, which should be prioritized for risk control, and the key factors affecting the safety of these pipe sections, are derived. Therefore, targeted measures can be taken on the identified pipe sections according to the key factors of each pipe section. The results are significant to help pipeline operators identify the key pipe sections and risk factors in the system and improve the efficiency of risk management.

3.2 Methodology

3.2.1 Risk identification

Identifying and classifying risk factors is an essential procedure in risk assessment, and

the Fishbone Diagram (FD) is an excellent way to carry it out. It is a method for discovering the root cause of a complex problem, which can also be called “Ishikawa” or “cause-and-branches” diagram [174], as shown in Figure 3.1. The process of risk analysis using fishbone diagrams is presented as follows. (1) Place the risk issue to be analyzed on the right side of the FD (i.e., at the head of the fish), and then draw the skeleton. (2) Identify the main categories of causes that generate risks and use them as the main branches. (3) Continue the investigation in depth based on the above categories, and expand each branch in layers until the underlying causes are analyzed [175].

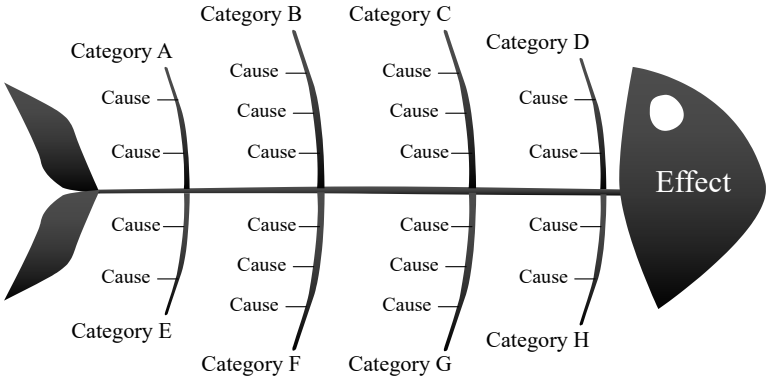


Figure 3.1 Basic structure of a FD Diagram

3.2.2 Risk assessment index system

A risk assessment index system can be established based on identifying risk factors [135]. The risk assessment index system is a hierarchical structure [115]. Figure 3.2 shows the three-layer construction of the index system. In Figure 3.2, the first-level index represents the risk status of the pipeline system. The second-level index is the main factor leading to system risk. The third-level index is subdivided from the second-level index, which is the basic cause of system risk.

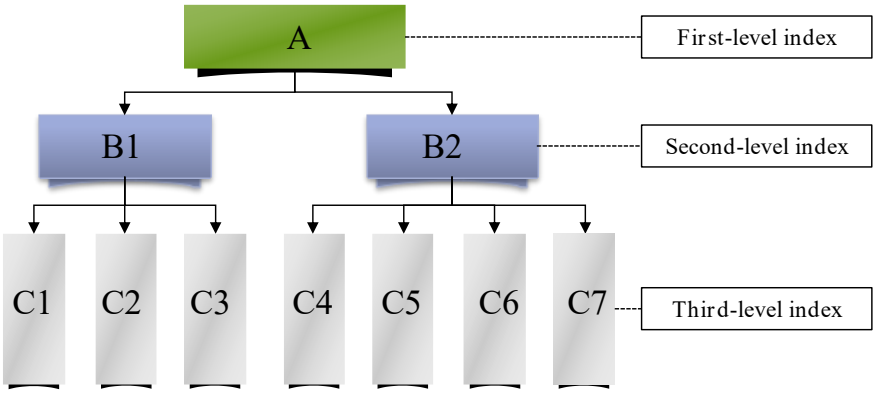


Figure 3.2 Three-layer construction of the index system

3.2.3 Cloud model theory

CM can be used to express the ambiguity and uncertainty of membership in a complex system, and it can convert a qualitative concept into a quantitative numerical representation. Fuzzy theory and probability statistics are the foundation of CM theory. However, according to traditional fuzzy theory [176], the membership function of any element is uniquely determined. But for CM, every cloud is composed of many cloud droplets, which is not a definite function curve. CM can integrate the uncertainty of the evaluation process through its numerical characteristics of expectation (*Ex*), entropy (*En*), and hyper entropy (*He*) [177].

In CM, *Ex* is the central value of the qualitative concept in the quantitative domain and can best represent the characteristics of a qualitative concept [178]. *En* is a measure of the uncertainty range of a qualitative concept and is determined by the fuzziness and randomness of the qualitative concept. It is represented as the opening degree of the cloud drop curve in Figure 3.3. *He* is a measure of uncertainty of *En* and it is expressed as the thickness of the cloud drop curve in Figure 3.3.

Forward cloud generator uses *Ex*, *En* and *He* to generate cloud drops. Based on the forward cloud generator, input the parameters of cloud characteristics and the number of cloud drops *N* (here, *N*=1000), and then a visual cloud model can be generated. The specific steps are shown as follows:

(1) Obtain the bounded interval [*Cmin*, *Cmax*] according to the actual research

problem. The formula for generating *Ex* [179]:

$$Ex = \left(C_{\max} + C_{\min}\right)/2 \tag{3.1}$$

(2) According to the "3 *En* Principle" of cloud theory [180], its main interval distribution is [*Ex*-3*En*, *Ex*+3*En*] for the qualitative concept *C* in the quantitative domain *U*. The cloud drops falling outside this interval can be ignored, and the formula for generating En can be obtained as:

$$En = \left(C_{\max} - C_{\min}\right)/6 \tag{3.2}$$

(3) Define He as a specific constant. To summarize, scholars usually use two common approaches to define the value, one of which is to establish a linear relationship between *He* and *En* (i.e., $He = k \times En$, see [181] for details). The second is to specify *He* as a constant based on expert estimations and practical situations. For example, in [179, 182, 183], *He* = 0.001, 0.005, 0.05, or 0.5. Because of the linear relationship, the former approach is strongly influenced by the value of *En*, and the uncertainty degree may not be controllable. Therefore, in order to keep the uncertainty present in the risk assessment process within a permissible level, the latter approach is adopted. According to the actual conditions, the *He* value is defined as a constant equal to 0.05 in this chapter, considered a low and allowable level of uncertainty.

(4) Calculate the membership function $\mu_C(x)$ with Eq. (3.3):

$$\mu_C\left(x\right) = e^{-\frac{\left(x_i - Ex\right)^2}{2\left(En_i\right)^2}} \tag{3.3}$$

Note that in CM, *U* is set as a quantitative domain, and *C* is the qualitative notion of *U*. Assuming the quantitative value $x \in U$. *x* is a random realization of the qualitative concept *C*, expressed as $\mu_C(x) \in [0,1]$. The specific expression is given as follows:

$$\mu_C(x): U \rightarrow [0, 1], \forall x \in U, x \rightarrow \mu_C(x). \tag{3.4}$$

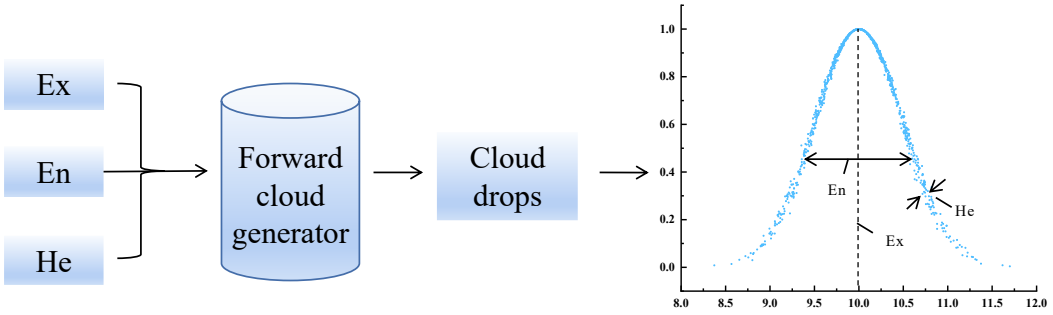


Figure 3.3 Steps to generate CM

3.2.4 Variable weight theory

Variable weight function mainly establishes the linkage between weight vector and state vector. It adjusts the weight of each indicator according to the change of the corresponding comment value (i,e. value of *Ex* of a comment cloud) to realize the reasonable distribution of weight in the process of risk assessment [184, 185]. In the risk assessment index system, sometimes an index has a smaller impact on the risk of the entire pipeline system than other indexes, and so when a constant weight is assigned, its value will be relatively small. However, indexes with small weights will also have great risks so that the comment value will be very low during expert evaluation. Because of the small weight, the influence of the index with greater risk on the final assessment result cannot be effectively reflected by the constant weight calculation method. For example, when the comment value of a second-level index is much smaller than the other comment values and its weight is relatively small, it will cause the misjudgment that the assessment result still seems promising when it is not that acceptable. Therefore, in this way, the actual risk status of the pipeline cannot be reflected precisely.

The steps for applying the VWT to solve a practical problem are shown as follows:

(1) Use variable weight function to construct the state-variable-weight vector. In this chapter, the zoning variable weight function [186] is used.

$$S_i(x)=\begin{cases}\dfrac{c_2-c_1}{\lambda-\mu}\mu\ln\dfrac{\mu}{x_i}+c_2 & 0<x_i\le\mu\\ -\dfrac{c_2-c_1}{\lambda-\mu}x_i+\dfrac{c_2\lambda-c_1\mu}{\lambda-\mu} & \mu<x_i\le\lambda\\ P+\dfrac{c_2-c_1}{2(\lambda-\mu)(\alpha-\lambda)}(\alpha-x_i)^2 & \lambda<x_i\le\alpha\\ P & \alpha<x_i\le\beta\\ Q(1-\beta)\ln\dfrac{1-\beta}{1-x_i}+P & \beta<x_i\le 1\end{cases} \tag{3.5}$$

where $S_i(x)$ represents the state-variable-weight vector, μ, λ, α and β represent the interval thresholds, x_i represents the comment value ith index, c_1, c_2, P and Q are constants determined by the actual condition.

(2) Determine the interval thresholds. The comment for each index needs to be normalized in this study.

(3) Calculate the constant weight of indexes. The cloud-AHP algorithm and virtual floating cloud computation are used for this step.

(4) Calculate the variable weight of indexes by Eq. (3.6) [187].

$$W_i(x)=\frac{w_i x_i}{\sum_{i=1}^{m} w_i x_i} \tag{3.6}$$

where $W_i(x)$ represents the variable weight of the ith index, i=1, 2, 3, …, m, m represents the number of the indexes, and w_i represents the constant weight of the ith index.

3.2.5 The proposed method

The specific steps of the proposed method are shown in Figure 3.4. A risk assessment index system is established based on the identified risk factors. The comment cloud matrix of the second-level indexes is generated by combining the constant weight cloud matrix and comment cloud matrix of third-level indexes. According to the generated comments and constant weight cloud matrix of the second-level indexes, the state-variable-weight vectors are calculated. Finally, the pipeline risk level is obtained.

According to Figure 3.4, it can be found that the calculation of CM runs through the entire evaluation process, including constant weight calculation, risk comment

model establishment, Delphi iteration, and variable weight calculation. Some existing methods only construct the comment cloud model and ignore the uncertainties in other steps [134, 159]. Some use fuzzy operators for evaluation, which can effectively express the uncertainty of the weight calculation process, but participate in the final result as a specific value after defuzzification, still losing part of the uncertainty [46, 188]. All these can make the results less reliable. In contrast, this chapter introduces the calculation of *En* and *He* in the whole process, and can always use *En* and *He* to express uncertainty. At the same time, the Delphi method can also optimize the expert comments and reduce human errors. According to the characteristics of the risk comments, the zoning variable weight function is used for calculation. Compared with existing methods [189, 190], the feedback optimization of constant weights based on the risk comment values is added, making the conclusion more reliable.

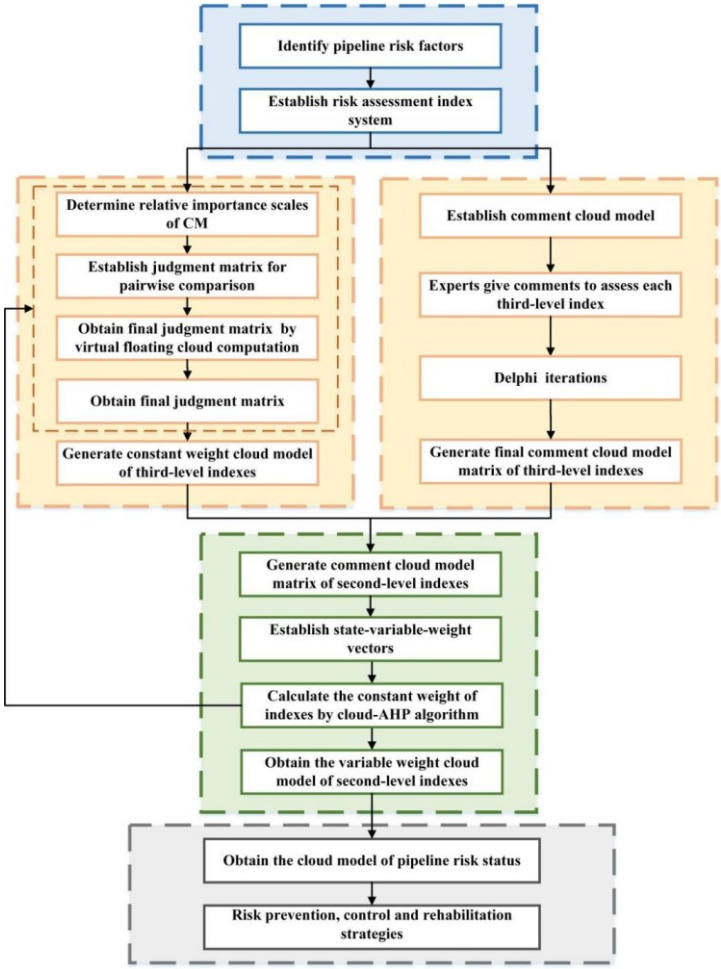


Figure 3.4 Schematic diagram of the proposed method

3.2.5.1 Comment cloud model

In this step, comment CMs are established, and experts give comments on the third-level indexes. The comment CMs transfer expert comments into CMs expressed by (*Ex*, *En*, *He*). There are three risk levels identified in the ALARP concept [191]: Level 1, Intolerable region; Level 2, ALARP region; Level 3, Tolerable region. Furthermore, to better manage the risks at different levels, many scholars divide the risks into five levels when conducting risk assessments [135, 192]. Therefore, in this chapter, the assessment results are divided into five levels to form a comment set *V*. Using Eq. (2) and fuzzy set theory in [193, 194], the *En* value can be calculated, with results shown in Table 3.2 and Figure 3.5.

Table 3.2 Numerical expression of comment CMs

Comment on the risk level	Numerical expression	Symbol	Explanation
High	(1,0.17,0.05)	*V1*	Risks must be reduced, measures must be taken until the risks are reduced to a tolerable level, and the reliability of emergency facilities should be ensured.
Relatively high	(2,0.33,0.05)	*V2*	In addition to enhanced management and monitoring, direct measures should be taken to perform risk rehabilitation.
Medium	(3,0.17,0.05)	*V3*	The effectiveness and reliability of monitoring and alarm facilities need to be strengthened, and high-quality management measures need to be taken on a cost-benefit basis.
Relatively low	(4,0.33,0.05)	*V4*	The risks are generally tolerable, and no extra measures are needed.
Low	(5,0.17,0.05)	*V5*	The risks are negligible, and no extra measures are needed.

There are *N* experts to participate in decision-making. *I* and *J* represent two different experts respectively. In the decision-making process, if opinions are inconsistent, it is necessary to use the Delphi method to carry out iterations, and gather all experts' assessment results to generate the final comment CM matrix, as shown in Eqs. (3.7) to (3.9) [190].

$$Ex_i = \frac{(Ex_i)_1+(Ex_i)_2+\cdots+(Ex_i)_N}{N} \tag{3.7}$$

$$He_i = \frac{(He_i)_1 + (He_i)_2 + \cdots + (He_i)_N}{N} \tag{3.8}$$

$$En_i = \frac{1}{6}\left[\max_I\{(Ex_i)_I + 3(En_i)_I\} - \min_J\{(Ex_i)_J - 3(En_i)_J\}\right] \tag{3.9}$$

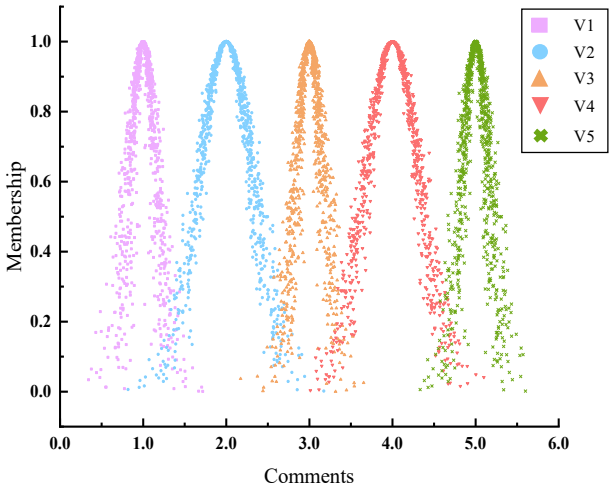


Figure 3.5 Image expression of comment CMs

The comprehensive assessment result can be obtained from the comment matrix and the weight matrix, which is calculated by:

$$C\left(Ex_i', En_i', He_i'\right) = \sum_{i=1}^{n} C\left(W_i(Ex_i), W_i(En_i), W_i(He_i)\right) \bullet C\left(Ex_i, En_i, He_i\right) \tag{3.10}$$

3.2.5.2 State-variable-weight vectors

In this section, cloud theory is used to improve the theory of variable weight function to make it more accurate and objective while expressing the uncertainty of assessment results.

Through the normalization, the operation rules of cloud [157] and Eq. (3.5), the state-variable-weight vectors of *Ex*, *En*, and *He* may be obtained.

$$Ex_{S_i(x)} = \begin{cases} \dfrac{c_2 - c_1}{\lambda - \mu}\mu \ln \dfrac{\mu}{x_i} + c_2 & 0 < x_i \le \mu \\ -\dfrac{c_2 - c_1}{\lambda - \mu}x_i + \dfrac{c_2\lambda - c_1\mu}{\lambda - \mu} & \mu < x_i \le \lambda \\ P + \dfrac{c_2 - c_1}{2(\lambda - \mu)(\alpha - \lambda)}(\alpha - x_i)^2 & \lambda < x_i \le \alpha \\ P & \alpha < x_i \le \beta \\ Q(1-\beta)\ln\dfrac{1-\beta}{1-x_i} + P & \beta < x_i \le 1 \end{cases}, \quad x_i = \frac{Ex_{C(Bi)}}{Ex_{C(V5)}} \tag{3.11}$$

$$En_{S_i(x)} = \sqrt{\frac{Ex_{S_i(x)} \bullet Ex_{C(Bi)}}{Ex_{C(V5)}}} \bullet \sqrt{\left(\frac{En_{C(Bi)}}{Ex_{C(Bi)}}\right)^2 + \left(\frac{En_{C(V5)}}{Ex_{C(V5)}}\right)^2} \tag{3.12}$$

$$He_{S_i(x)} = \sqrt{\frac{Ex_{S_i(x)} \bullet Ex_{C(Bi)}}{Ex_{C(V5)}}} \bullet \sqrt{\left(\frac{He_{C(Bi)}}{Ex_{C(Bi)}}\right)^2 + \left(\frac{He_{C(V5)}}{Ex_{C(V5)}}\right)^2} \tag{3.13}$$

where C(*Bi*) represents the cloud model of the second-level index, and C(*V5*) represents the cloud model with a comment level of *V5*.

3.2.5.3 Calculation of constant weight based on Cloud-AHP

In the index system, the index weights of the second and third levels can be derived from the experience and knowledge of experts and the calculation of CMs. It is necessary to take the error caused by the personal preference of the experts into consideration. Therefore, the virtual floating cloud computation is introduced to integrate the weight values given by each expert to generate the final weight value. This research proposes an improved cloud-AHP algorithm to quantify the relative importance of each index.

(1) The CM is integrated into the traditional AHP and a relative importance scale of CM is built. Experts use this scale to measure the weight of each index. Based on Eqs. (3.1) and (3.2) and the scale of traditional AHP, the relative importance scales of the nine CMs are defined as shown in Table 3.3.

(2) Establish a judgment matrix for pairwise comparison from each expert:

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} C_{11}(Ex_{11}, En_{11}, He_{11}) & C_{12}(Ex_{12}, En_{12}, He_{12}) & \cdots & C_{1n}(Ex_{1n}, En_{1n}, He_{1n}) \\ C_{21}(Ex_{21}, En_{21}, He_{21}) & C_{22}(Ex_{22}, En_{22}, He_{22}) & \cdots & C_{2n}(Ex_{2n}, En_{2n}, He_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1}(Ex_{n1}, En_{n1}, He_{n1}) & C_{n2}(Ex_{n2}, En_{n2}, He_{n2}) & \cdots & C_{nn}(Ex_{nn}, En_{nn}, He_{nn}) \end{pmatrix} \tag{3.14}$$

(3) Perform integration operations using virtual floating cloud computation [134] on the above judgment matrix, as shown in Eqs. (3.15) to (3.17).

$$Ex_{ij} = \alpha_1 \left(Ex_{ij}\right)_1 + \alpha_2 \left(Ex_{ij}\right)_2 + \cdots + \alpha_N \left(Ex_{ij}\right)_N \tag{3.15}$$

$$En_{ij} = \frac{\alpha_1^{\,2}}{\alpha_1^{\,2} + \alpha_2^{\,2} + \cdots + \alpha_N^{\,2}} \left(En_{ij}\right)_1 + \frac{\alpha_2^{\,2}}{\alpha_1^{\,2} + \alpha_2^{\,2} + \cdots + \alpha_N^{\,2}} \left(En_{ij}\right)_2 + \cdots + \frac{\alpha_N^{\,2}}{\alpha_1^{\,2} + \alpha_2^{\,2} + \cdots + \alpha_N^{\,2}} \left(En_{ij}\right)_N \tag{3.16}$$

Table 3.3 Relative importance scales of CM [52]

Relative importance a_{ij}	CM scale	Description
1	C_1=(1,0,0)	i and j are equally important
3	C_3=(3,0.33,0.05)	i is slightly more important than j
5	C_5=(5,0.33,0.05)	i is obviously more important than j
7	C_7=(7,0.33,0.05)	i is strongly more important than j
9	C_9=(9,0.33,0.05)	i is absolutely more important than j
1/3	$C_{1/3}$=(1/3,0.33/9,0.05/9)	j is slightly more important than i
1/5	$C_{1/5}$=(1/5,0.33/25,0.05/25)	j is obviously more important than i
1/7	$C_{1/7}$=(1/7,0.33/49,0.05/49)	j is strongly more important than i
1/9	$C_{1/9}$=(1/9,0.33/81,0.05/81)	j is absolutely more important than i

$$He_{ij}=\frac{\alpha_1^{\,2}}{\alpha_1^{\,2}+\alpha_2^{\,2}+\cdots+\alpha_N^{\,2}}\left(He_{ij}\right)_1+\frac{\alpha_2^{\,2}}{\alpha_1^{\,2}+\alpha_2^{\,2}+\cdots+\alpha_N^{\,2}}\left(He_{ij}\right)_2+\cdots+\frac{\alpha_N^{\,2}}{\alpha_1^{\,2}+\alpha_2^{\,2}+\cdots+\alpha_N^{\,2}}\left(He_{ij}\right)_N \tag{3.17}$$

where α=(α_1,α_2,α_3,…,α_N) is the weight vector, representing the relative importance of a specific expert to all experts. (Here, experts have the same experience, knowledge, and education, so α_1=α_2=α_3=1/3.) N represents the number of experts, N=1,2,3…

(4) According to the traditional AHP and the operation rules of CM, the product of each row element of the judgment matrix is calculated as follows:

$$Ex_i=\prod_{j=1}^{n}Ex_{ij} \tag{3.18}$$

$$En_i=\prod_{j=1}^{n}Ex_{ij}\sqrt{\sum_{j=1}^{n}\left(\frac{En_{ij}}{Ex_{ij}}\right)^2} \tag{3.19}$$

$$He_i=\prod_{j=1}^{n}Ex_{ij}\sqrt{\sum_{j=1}^{n}\left(\frac{He_{ij}}{Ex_{ij}}\right)^2} \tag{3.20}$$

(5) Perform normalization calculation, and obtain the index weight as:

$$w\left(Ex_i\right)=\frac{\sqrt[n]{Ex_i}}{\sum_{i=1}^{n}\sqrt[n]{Ex_i}} \tag{3.21}$$

$$w\left(En_i\right)=\frac{\sqrt[n]{Ex_i}}{\sum_{i=1}^{n}\sqrt[n]{Ex_i}}\bullet\sqrt{\left(\frac{\sqrt{\frac{1}{n}}\bullet\left(Ex_i\right)^{\frac{1}{n}-1}\bullet En_i}{\sqrt[n]{Ex_i}}\right)^2+\left(\frac{\sum_{i=1}^{n}\sqrt{\frac{1}{n}}\bullet\left(Ex_i\right)^{\frac{1}{n}-1}\bullet En_i}{\sum_{i=1}^{n}\sqrt[n]{Ex_i}}\right)^2} \tag{3.22}$$

$$w(He_i)=\frac{\sqrt[n]{Ex_i}}{\sum_{i=1}^{n}\sqrt[n]{Ex_i}}\bullet\sqrt{\left(\frac{\sqrt{\frac{1}{n}}\bullet(Ex_i)^{\frac{1}{n}-1}\bullet He_i}{\sqrt[n]{Ex_i}}\right)^2+\left(\frac{\sum_{i=1}^{n}\sqrt{\frac{1}{n}}\bullet(Ex_i)^{\frac{1}{n}-1}\bullet He_i}{\sum_{i=1}^{n}\sqrt[n]{Ex_i}}\right)^2} \tag{3.23}$$

3.2.5.4 Calculation of the variable weight

According to Eq. (3.6) and Eqs. (3.11) to (3.13), the variable weight of *Ex*, *En*, and *He* can be calculated by Eqs. (3.24) to (3.26).

$$W_{Bi}(Ex)=\frac{Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}}{\sum_{i=1}^{n}Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}} \tag{3.24}$$

$$W_{Bi}(En)=\frac{Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}}{\sum_{i=1}^{n}Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}}\sqrt{(\frac{En_{S_i(x)}}{Ex_{S_i(x)}})^2+(\frac{En_{w_{C(Bi)}}}{Ex_{w_{C(Bi)}}})^2+\frac{\sum_{i=1}^{n}\left[Ex_{S_i(x)}\right]^2\bullet\left[Ex_{w_{C(Bi)}}\right]^2\bullet\left[(\frac{En_{S_i(x)}}{Ex_{S_i(x)}})^2+(\frac{En_{w_{C(Bi)}}}{Ex_{w_{C(Bi)}}})^2\right]}{\left(\sum_{i=1}^{n}Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}\right)^2}} \tag{3.25}$$

$$W_{Bi}(He)=\frac{Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}}{\sum_{i=1}^{n}Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}}\sqrt{(\frac{He_{S_i(x)}}{Ex_{S_i(x)}})^2+(\frac{He_{w_{C(Bi)}}}{Ex_{w_{C(Bi)}}})^2+\frac{\sum_{i=1}^{n}\left[Ex_{S_i(x)}\right]^2\bullet\left[Ex_{w_{C(Bi)}}\right]^2\bullet\left[(\frac{He_{S_i(x)}}{Ex_{S_i(x)}})^2+(\frac{He_{w_{C(Bi)}}}{Ex_{w_{C(Bi)}}})^2\right]}{\left(\sum_{i=1}^{n}Ex_{S_i(x)}\bullet Ex_{w_{C(Bi)}}\right)^2}} \tag{3.26}$$

where $w_{C(B_i)}$ represents the constant weight of the second-level index.

3.3 Case Study

3.3.1 Construction of the Fishbone Diagram

An industrial case study is presented to illustrate various steps of the proposed methodology. A section of a gas pipeline located in Beijing, China, is used as an example. On the left side of this pipe section is a subway line, and on the upper right side of the pipe section is a highway with many vehicles. According to the records, the pipe body is corroded. The basic information has been collected to identify and classify the risk factors. According to the procedure of FD, there can be five categories, as shown in Figure 3.6.

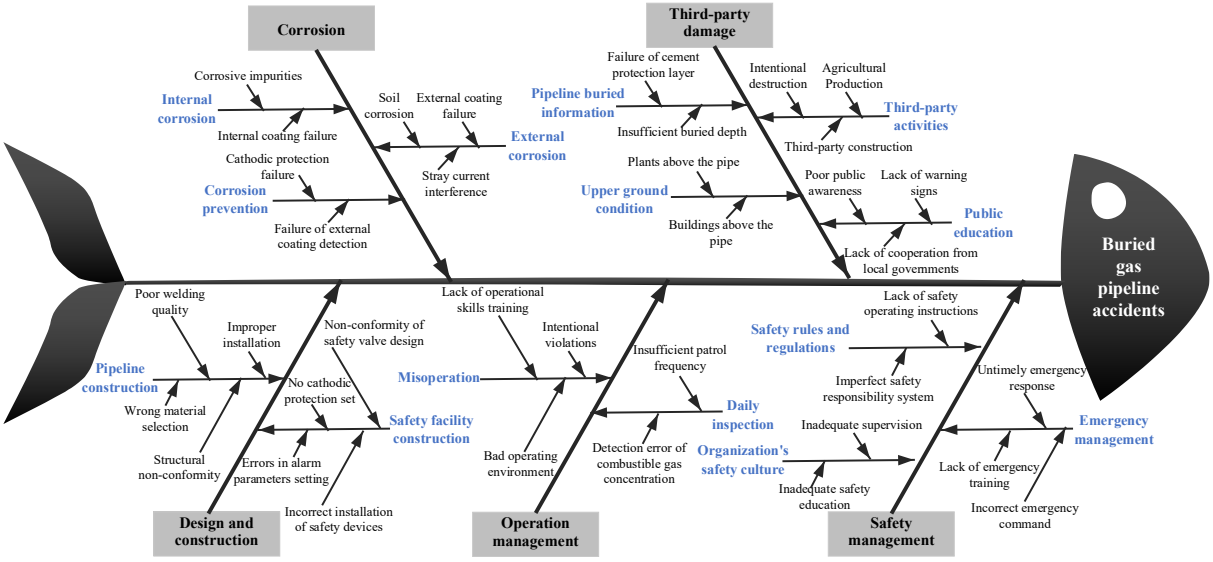


Figure 3.6 FD of the pipe section

3.3.2 Construction of the index system

Based on the risk identification and the hierarchical structure, we also added the consequence severity index to establish the risk index system, as shown in Figure 3.7. The consequence index is also a key factor in determining the priority of risk mitigation. C15 is used to measure the accident's impact on the external environment, including natural environment damage and economic loss. C16 mainly includes the population density and number of casualties that may result.

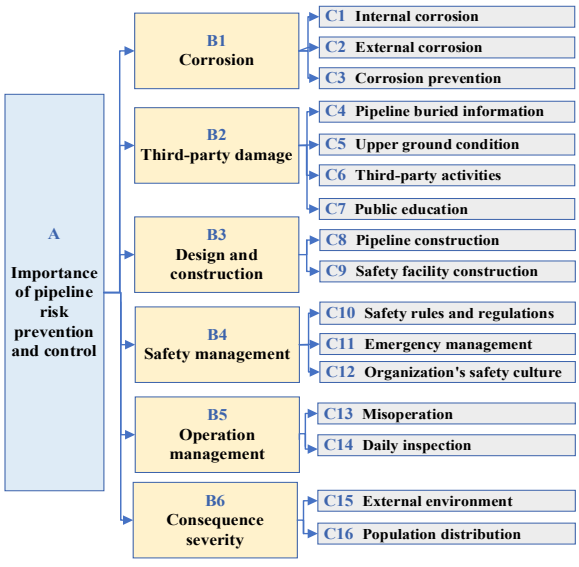


Figure 3.7 Risk assessment index system

3.3.3 Risk analysis based on the proposed method

3.3.3.1 Constant weight calculation based on weighted average cloud algorithm

(1) Weight calculation of third-level indexes

According to the index system established in Figure 3.7, each second-level index (i.e., Bi, i=1, 2,..., 6) has several branches of the third-level index (i.e., Cj, j=1,..., 4). Therefore, for the convenience of description, each second-level index and its branches are defined as Bi-C, and each matrix calculation is performed only for the third-level index branches of a second-level index. In this section, calculations are carried out for the B1-C matrix, B2-C matrix, B3-C matrix, B4-C matrix, B5-C matrix, and B6-C matrix. Taking the B2-C matrix as an example, the calculation process is presented.

According to Eq. (3.14), the judgment matrix for pairwise comparison given by three experienced experts (name them a, b, c) is shown in Tables 3.4 to 3.6.

Table 3.4 Judgment matrix of B2-C given by expert a

	C4	C5	C6	C7
C4	(1,0,0)	(1/5,0.33/25,0.05/25)	(1/7,0.33/49,0.05/49)	(1,0,0)
C5	(5,0.33,0.05)	(1,0,0)	(1/3,0.33/9,0.05/9)	(5,0.33,0.05)
C6	(7,0.33,0.05)	(3,0.33,0.05)	(1,0,0)	(7,0.33,0.05)
C7	(1,0,0)	(1/5,0.33/25,0.05/25)	(1/7,0.33/49,0.05/49)	(1,0,0)

Table 3.5 Judgment matrix of B2-C given by expert b

	C4	C5	C6	C7
C4	(1,0,0)	(1/5,0.33/25,0.05/25)	(1/5,0.33/25,0.05/25)	(3,0.33,0.05)
C5	(5,0.33,0.05)	(1,0,0)	(1,0,0)	(7,0.33,0.05)
C6	(5,0.33,0.05)	(1,0,0)	(1,0,0)	(5,0.33,0.05)
C7	(1/3,0.33/9, 0.05/9)	(1/7,0.33/49,0.05/49)	(1/5,0.33/25,0.05/25)	(1,0,0)

Table 3.6 Judgment matrix of B2-C given by expert c

	C4	C5	C6	C7
C4	(1,0,0)	(1/3,0.33/9,0.05/9)	(1/5,0.33/25,0.05/25)	(1/3,0.33/9,0.05/9)
C5	(3,0.33,0.05)	(1,0,0)	(1/3,0.33/9,0.05/9)	(3,0.33,0.05)
C6	(5,0.33,0.05)	(3,0.33,0.05)	(1,0,0)	(5,0.33,0.05)
C7	(3,0.33,0.05)	(1/3,0.33/9,0.05/9)	(1/5,0.33/25,0.05/25)	(1,0,0)

According to Eqs.(3.15) to (3.17), the final judgment matrix is obtained as shown in Table 8.

Table 3.7 Final judgment matrix of B2-C

	C4	C5	C6	C7
C4	(1,0,0)	(0.2444,0.0137, 0.0021)	(0.1810,0.0066, 0.0010)	(1.4444,0.1107, 0.0168)
C5	(4.3333,0.1905, 0.0289)	(1,0,0)	(0.5556,0.0173, 0.0026)	(5,0.1905,0.0289)
C6	(5.6667,0.1905, 0.0289)	(2.3333,0.1556, 0.0236)	(1,0,0)	(5.6667,0.1905, 0.0289)
C7	(1.4444,0.1107, 0.0168)	(0.2254,0.0131, 0.0020)	(0.1810,0.0066, 0.0010)	(1,0,0)

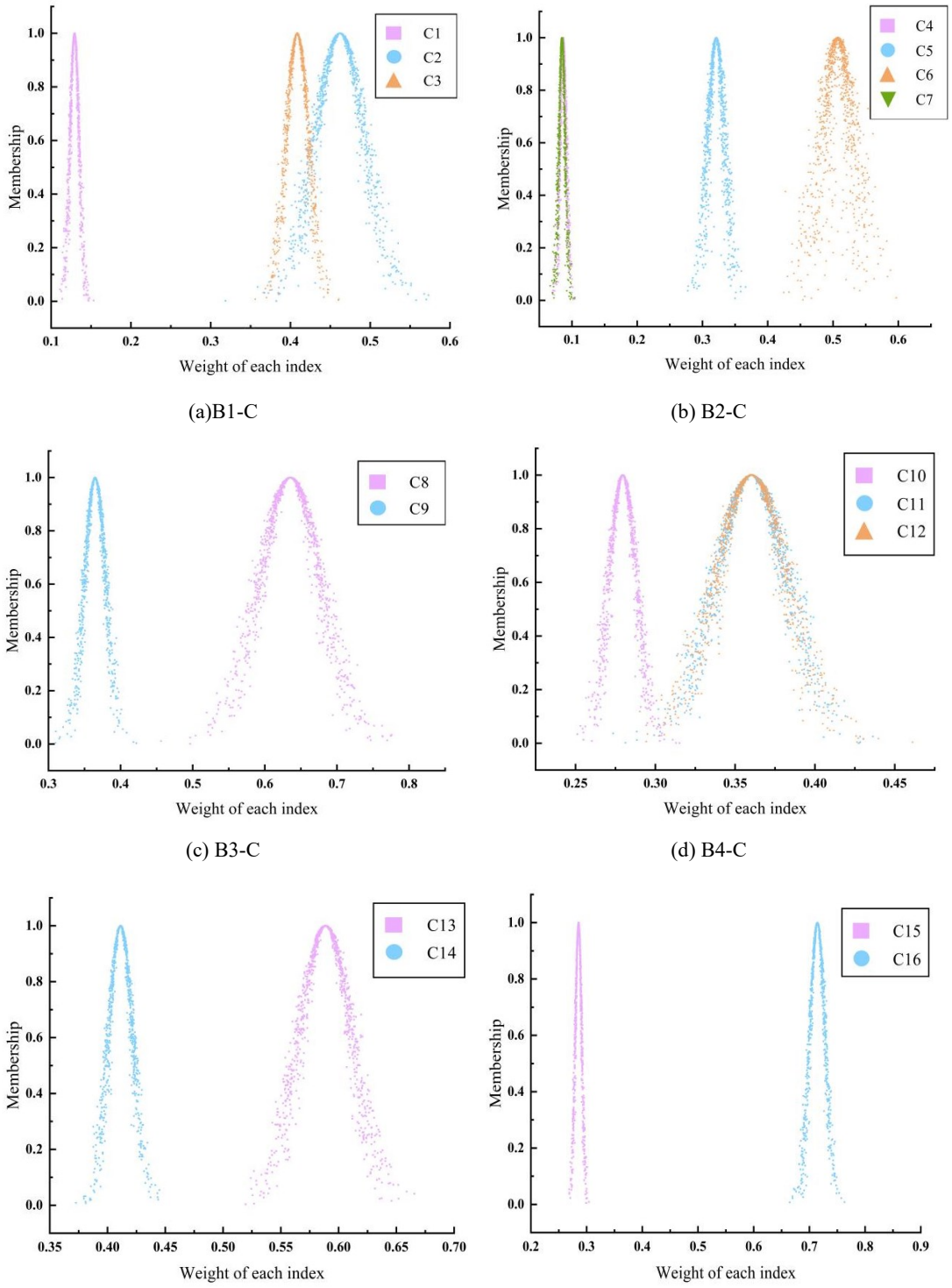
The numerical expressions of the weight cloud model of the third-level indexes C4, C5, C6, and C7 are calculated by Eqs. (3.21) to (3.23).

Similarly, the numerical expression of the weight CM of other third-level indexes can be obtained. The final calculation results are shown in

Table 3.8. Figure 3.8 shows the relative importance of each index more clearly through the expression of images.

Table 3.8 Numerical expression of the weight CM (third-level indexes)

Second-level indexes	Third-level indexes
B1	C1(0.1293, 0.0058, 0.0009)
	C2(0.4622, 0.0314, 0.0048)
	C3(0.4085, 0.0143, 0.0022)
B2	C4(0.0867, 0.0049, 0.0010)
	C5(0.3211, 0.0131, 0.0034)
	C6(0.5072, 0.0241, 0.0101)
B3	C7(0.0849, 0.0048, 0.0010)
	C8(0.6353, 0.0432, 0.0066)
	C9(0.3647, 0.0150, 0.0023)
B4	C10(0.2796, 0.0088, 0.0013)
	C11(0.3602, 0.0231, 0.0035)
	C12(0.3602, 0.0231, 0.0035)
B5	C13(0.5887, 0.0231, 0.0035)
	C14(0.4113, 0.0106, 0.0016)
B6	C15(0.2856, 0.0058, 0.0009)
	C16(0.7144, 0.0136, 0.0021)



(a)B1-C (b) B2-C

(c) B3-C (d) B4-C

(e) B5-C (f) B6-C

Figure 3.8 Image expression of the weight CM (third-level indexes)

(2) Weight calculation of second-level indexes

Similar to the calculation method of the B-C matrix, the calculation process of the A-B matrix is as follows. According to Eq. (3.14), the judgment matrix for pairwise

comparison given by three experienced experts (named as a, b, c) is shown in Tables 3.9 to 3.11.

Table 3.9 Judgment matrix of A-B given by expert a

	B1	B2	B3	B4	B5	B6
B1	(1,0,0)	(1,0,0)	(5,0.33,0.05)	(7,0.33,0.05)	(5,0.33,0.05)	(3,0.33,0.05)
B2	(1,0,0)	(1,0,0)	(5,0.33,0.05)	(5,0.33,0.05)	(7,0.33,0.05)	(5,0.33,0.05)
B3	(1/5,0.33/25, 0.05/25)	(1/5,0.33/25, 0.05/25)	(1,0,0)	(1/3,0.33/9, 0.05/9)	(1,0,0)	(1/5,0.33/25, 0.05/25)
B4	(1/7,0.33/49, 0.05/49)	(1/5,0.33/25, 0.05/25)	(3,0.33,0.05)	(1,0,0)	(3,0.33,0.05)	(1/3,0.33/9, 0.05/9)
B5	(1/5,0.33/25, 0.05/25)	(1/7,0.33/49, 0.05/49)	(1,0,0)	(1/3,0.33/9,0.05/ 9)	(1,0,0)	(1/3,0.33/9, 0.05/9)
B6	(1/3,0.33/9, 0.05/9)	(1/5,0.33/25, 0.05/25)	(5,0.33,0.05)	(3,0.33,0.05)	(3,0.33,0.05)	(1,0,0)

Table 3.10 Judgment matrix of A-B given by expert b

	B1	B2	B3	B4	B5	B6
B1	(1,0,0)	(1,0,0)	(3,0.33,0.05)	(9,0.33,0.05)	(5,0.33,0.05)	(3,0.33,0.05)
B2	(1,0,0)	(1,0,0)	(5,0.33,0.05)	(3,0.33,0.05)	(7,0.33,0.05)	(5,0.33,0.05)
B3	(1/3,0.33/9, 0.05/9)	(1/5,0.33/25, 0.05/25)	(1,0,0)	(1/3,0.33/9, 0.05/9)	(1,0,0)	(1/3,0.33/9, 0.05/9)
B4	(1/9,0.33/81, 0.05/81)	(1/3,0.33/9, 0.05/9)	(3,0.33,0.05)	(1,0,0)	(3,0.33,0.05)	(1/3,0.33/9, 0.05/9)
B5	(1/5,0.33/25, 0.05/25)	(1/7,0.33/49, 0.05/49)	(1,0,0)	(1/3,0.33/9, 0.05/9)	(1,0,0)	(1/3,0.33/9, 0.05/9)
B6	(1/3,0.33/9, 0.05/9)	(1/5,0.33/25, 0.05/25)	(3,0.33,0.05)	(3,0.33,0.05)	(3,0.33,0.05)	(1,0,0)

Table 3.11 Judgment matrix of A-B given by expert c

	B1	B2	B3	B4	B5	B6
B1	(1,0,0)	(3,0.33,0.05)	(5,0.33,0.05)	(7,0.33,0.05)	(5,0.33,0.05)	(3,0.33,0.05)
B2	(1/3,0.33/9, 0.05/9)	(1,0,0)	(3,0.33,0.05)	(3,0.33,0.05)	(5,0.33,0.05)	(5,0.33,0.05)
B3	(1/5,0.33/25, 0.05/25)	(1/3,0.33/9, 0.05/9)	(1,0,0)	(1,0,0)	(1,0,0)	(1/5,0.33/25, 0.05/25)
B4	(1/7,0.33/49, 0.05/49)	(1/3,0.33/9, 0.05/9)	(1,0,0)	(1,0,0)	(3,0.33,0.05)	(1/3,0.33/9, 0.05/9)
B5	(1/5,0.33/25, 0.05/25)	(1/5,0.33/25, 0.05/25)	(1,0,0)	(1/3,0.33/9,0.05/ 9)	(1,0,0)	(1/3,0.33/9, 0.05/9)
B6	(1/3,0.33/9, 0.05/9)	(1/5,0.33/25, 0.05/25)	(5,0.33,0.05)	(3,0.33,0.05)	(3,0.33,0.05)	(1,0,0)

According to Eqs. (3.15) to (3.17), the final judgment matrix is obtained as shown

in Table 3.12.

Table 3.12 Final judgment matrix of A-B

	B1	B2	B3	B4	B5	B6
B1	(1,0,0)	(1.6667,0.1100, 0.0167)	(4.3333,0.1905, 0.0289)	(7.6667,0.1905, 0.0289)	(5,0.1905,0.0289)	(3,0.1905, 0.0289)
B2	(0.7778,0.03 67,0.0056)	(1,0,0)	(4.3333,0.1905, 0.0289)	(2.2,0.1905,0.0 289)	(6.3333,0.1905,0. 0289)	(5,0.1905, 0.0289)
B3	(0.2444,0.01 37,0.0021)	(0.2444,0.0137, 0.0021)	(1,0,0)	(0.5556,0.0173, 0.0026)	(1,0,0)	(0.2444,0.0137, 0.0021)
B4	(0.1323,0.00 35,0.0005)	(0.2889,0.0178, 0.0027)	(2.3333,0.1556, 0.0236)	(1,0,0)	(3,0.1905,0.0289)	(0.3333,0.0212, 0.0032)
B5	(0.2,0.0076, 0.0012)	(0.1619,0.0054, 0.0008)	(1,0,0)	(0.3333,0.0212, 0.0032)	(1,0,0)	(0.3333,0.0212, 0.0032)
B6	(0.3333,0.02 12,0.0032)	(0.2,0.0076, 0.0012)	(4.3333,0.1905, 0.0289)	(3,0.1905,0.028 9)	(3,0.1905,0.0289)	(1,0,0)

The numerical expressions of the weight CM of the second-level indexes (Table 3.13) are calculated by Eqs. (3.21) to (3.23). Figure 3.9 shows the relative importance of each index more clearly through the expression of images.

Table 3.13 Numerical expression of the weight CM (second-level indexes)

Second-level indexes	Numerical expressions	Rank
B1	(0.3725, 0.0192, 0.0040)	1
B2	(0.3018, 0.0163, 0.0033)	2
B3	(0.0545, 0.0026, 0.0008)	5
B4	(0.0812, 0.0047, 0.0009)	4
B5	(0.0476, 0.0023, 0.0013)	6
B6	(0.1425, 0.0080, 0.0040)	3

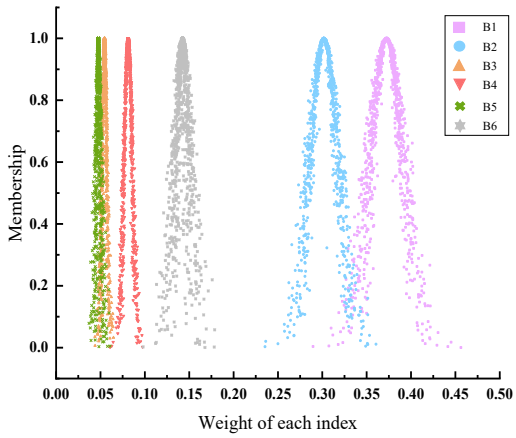


Figure 3.9 Image expression of the weight CM (second-level indexes)

3.3.3.2 Comment calculation of third-level indexes

Take the B2-C matrix as an example. Several experts' comments on C4, C5, C6, and C7 are shown in Tables 3.14 to 3.16 below.

Table 3.14 Comments on B2-C from expert a

C4	C5	C6	C7
*V*4(4,0.33,0.05)	*V*1(1,0.17,0.05)	*V*2(2,0.33,0.05)	*V*2(2,0.33,0.05)

Table 3.15 Comments on B2-C from expert b

C4	C5	C6	C7
*V*4(4,0.33,0.05)	*V*2(2,0.33,0.05)	*V*3(3,0.17,0.05)	*V*4(4,0.33,0.05)

Table 3.16 Comments on B2-C from expert c

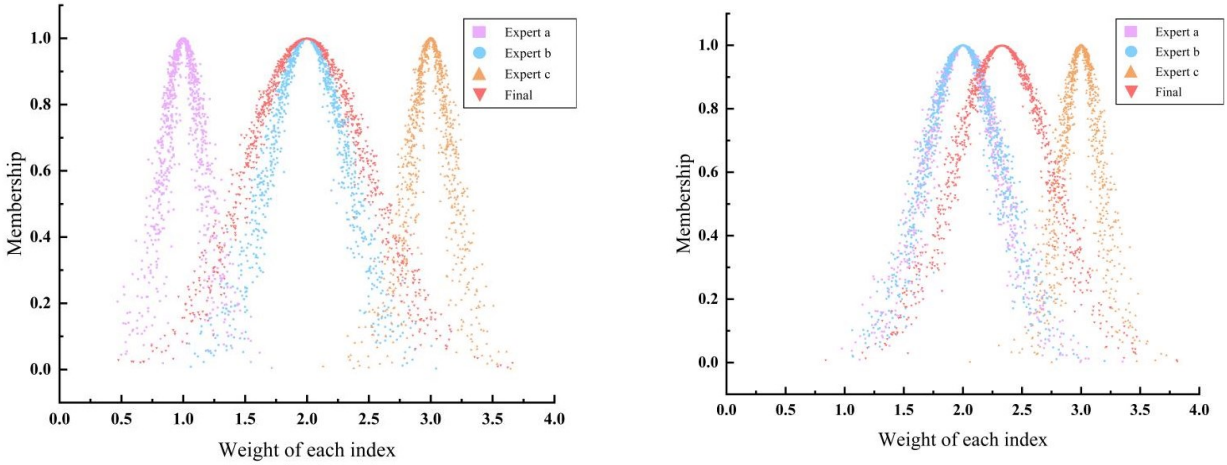
C4	C5	C6	C7
*V*5(5,0.17,0.05)	*V*3(3,0.17,0.05)	*V*2(2,0.33,0.05)	*V*5(5,0.17,0.05)

The assessment results of all experts are gathered using Eqs.(3.7) to (3.9) as shown in Table 3.17.

Table 3.17 Final comments on B2-C

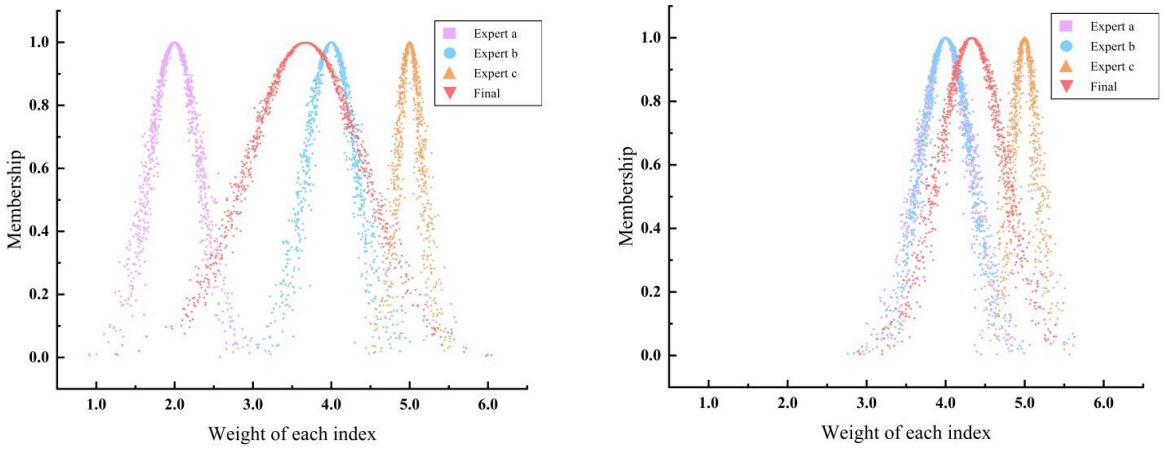
C4	C5	C6	C7
(4.33,0.42,0.05)	(2,0.50,0.05)	(2.33,0.42,0.05)	(3.67,0.75,0.05)

The comparison reveals a disagreement in the experts' opinion in the evaluations for C5 and C7. Therefore, the Delphi method is used to carry out iterative feedback to collect further information related to C5 and C7. The re-evaluations should continue until the evaluation consistency meets the requirements. The comments after iterative feedback are shown in Tables 3.18 to 3.21. In Figures 3.10 and 3.11, it is clear that the aggregation and consistency of the experts' comments are higher after completing iterative feedback using the Delphi method.



(a) (b)

Figure 3.10 Comparison of C5 before and after iterative feedback



(a) (b)

Figure 3.11 Comparison of C7 before and after iterative feedback

Table 3.18 Comments from expert a on B2-C after iterative feedback

C4	C5	C6	C7
*V*4(4,0.33,0.05)	*V*2(2,0.33,0.05)	*V*2(2,0.33,0.05)	*V*4(4,0.33,0.05)

Table 3.19 Comments from expert b on B2-C after iterative feedback

C4	C5	C6	C7
*V*4(4,0.33,0.05)	*V*2(2,0.33,0.05)	*V*3(3,0.17,0.05)	*V*4(4,0.33,0.05)

Table 3.20 Comments from expert c on B2-C after iterative feedback

C4	C5	C6	C7
*V*5(5,0.17,0.05)	*V*3(3,0.17,0.05)	*V*2(2,0.33,0.05)	*V*5(5,0.17,0.05)

Table 3.21 Final comments on B2-C after iterative feedback

C4	C5	C6	C7
(4.33,0.42,0.05)	(2.33,0.42,0.05)	(2.33,0.42,0.05)	(4.33,0.42,0.05)

According to Eq. (3.10), the numerical expression of the cloud model for the final comment of B2 can be obtained as C(B2)=(2.673, 0.2667, 0.0399). Similarly, the final comments on the other second-level indexes are shown in Table 3.22.

Table 3.22 Numerical expression of the comment cloud model (second-level indexes)

Second-level indexes	Numerical expressions
B1	(1.4445, 0.2660,0 .0327)
B2	(2.6730, 0.2667, 0.0399)
B3	(3.0594, 0.3302, 0.0410)
B4	(3.1612, 0.2555, 0.0352)
B5	(3.9415, 0.3167, 0.0388)
B6	(1.1914, 0.1715, 0.0386)

3.3.3.3 Variable weight calculation of the second-level indexes

According to Eq. (3.11) and Table 3.2, the value of the constants is obtained after normalization calculation. The state-variable-weight vector of *Ex* is constructed as shown in Eq. (3.27).

$$Ex_{S_i(x)}=\begin{cases}0.2\ln\dfrac{0.2}{x_i}+0.4 & 0<x_i\le 0.2\\ -x_i+0.6 & 0.2<x_i\le 0.4\\ 0.1+2.5\left(0.6-x_i\right)^2 & 0.4<x_i\le 0.6,\\ 0.1 & 0.6<x_i\le 0.8\\ 0.2\ln\dfrac{0.2}{1-x_i}+0.1 & 0.8<x_i\le 1\end{cases}\quad x_i=\frac{Ex_{C(B_i)}}{Ex_{C(V5)}} \tag{3.27}$$

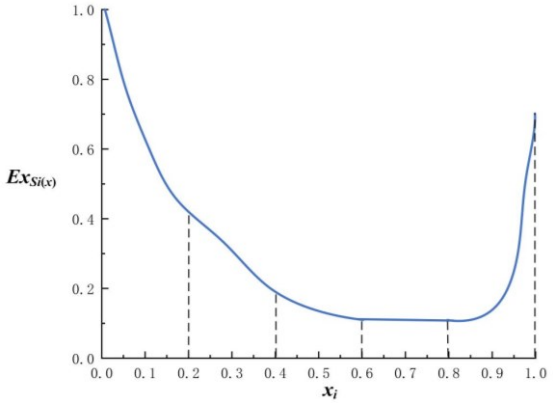


Figure 3.12 Image expression of the variable weight function

As shown in Figure 3.12, the function is continuous. When $0< x_i \leq 0.6$, the function image shows a downward trend, which tends to "punish" the risk level. When $0.6< x_i \leq 0.8$, the function's slope is zero, indicating that there is neither punishment nor incentive on this interval. When $0.8< x_i \leq 1$, the function image shows an upward trend, which tends to "reward" the risk level. Besides, when $0< x_i \leq 0.2$, the function slope is much higher, indicating that the heaviest penalty should be imposed by increasing the index's weight due to the high pipeline risk. When $0.8< x_i \leq 1$, The rapid rise of the function indicates that it is necessary to increase the index's weight to give reasonable rewards to the pipeline with very low risk.

The state-variable-weight vectors of *Ex*, *En*, *He* and the variable weight of *Ex*, *En*, *He* are calculated by Eqs. (3.24) to (3.26), as shown in Table 3.23.

Table 3.23 The state-variable-weight vectors and the variable weight

Second-level indexes	State-variable-weight vectors	Variable weight	Rank
B1	(0.3111, 0.0561, 0.0074)	(0.5288, 0.1152, 0.0167)	1
B2	(0.1107, 0.0256, 0.0044)	(0.1524, 0.0400, 0.0068)	3
B3	(0.1000, 0.0280, 0.0041)	(0.0249, 0.0076, 0.0012)	5
B4	(0.1000, 0.0220, 0.0037)	(0.0370, 0.0094, 0.0016)	4
B5	(0.1000, 0.0245, 0.0039)	(0.0217, 0.0059, 0.0011)	6
B6	(0.3617, 0.0434, 0.0095)	(0.2352, 0.0406, 0.0100)	2

After substituting the variable weight and the comments of the second-level index into Eq. (3.10), the final assessment result is obtained as (1.7301,0.2583,0.0392) as shown in Figure 3.13.

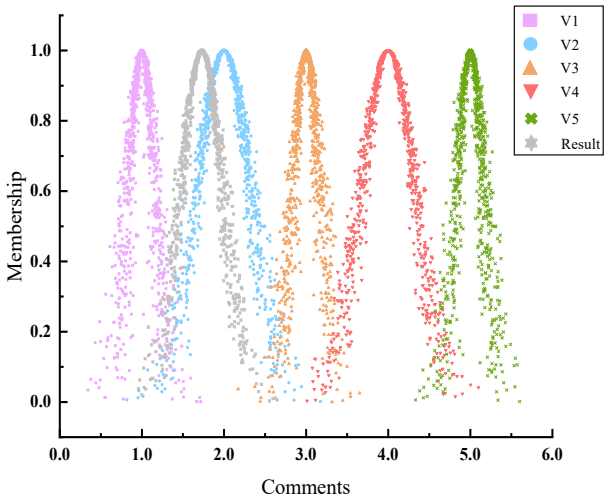


Figure 3.13 Final assessment result

3.4 Discussions

3.4.1 Assessment of overall risk

This chapter establishes a risk assessment index system for gas pipelines. It optimizes the importance of each second-level index for the risk of gas pipelines by using cloud-variable weight theory. In order to verify the correctness of the method proposed in this chapter, we use a risk assessment method based on triangular fuzzy numbers [46] to evaluate the research object in this chapter. The result is 1.8383, similar to the *Ex* value of the result calculated by the proposed method, which is defined as "relatively high risk." Therefore, the proposed method in this chapter is feasible and applicable. In addition, the proposed method calculates the result as a cloud instead of an exact number, thus enriching the dimensionality of the result expression, which better expresses the uncertainty inherent to the entire evaluation process based on the characteristics of the cloud model.

3.4.2 Sensitivity Analysis

Sensitivity analysis is used to assess the impact of different sub-indexes on the overall risk. The resulting ranking of index sensitivities will make a valuable contribution to decisions on risk prevention and control. A given index's sensitivity is determined by the difference in the final risk assessment result as its level gradually changes from *V1* to *V5* [154]. In this process, take index C1 as an example. C1 is considered one of the five levels, while the other indexes' levels are randomly generated from level *V1* to level *V5*. To guarantee the robustness of the outcomes, 1000 simulations are performed for each level of the specified index C1. The final result *R* is calculated by the average of the outcomes of the 1000 simulations. The sensitivity of the third-level indexes can be calculated by Eq. (28). As shown in Figure 3.14, the *Ex* value of the final assessment result increases, i.e. the risk decreases, as the level of indexes increases from level *V1* to level *V5*. To ensure clarity, only 10 out of 16 third-level indexes are shown in Figure

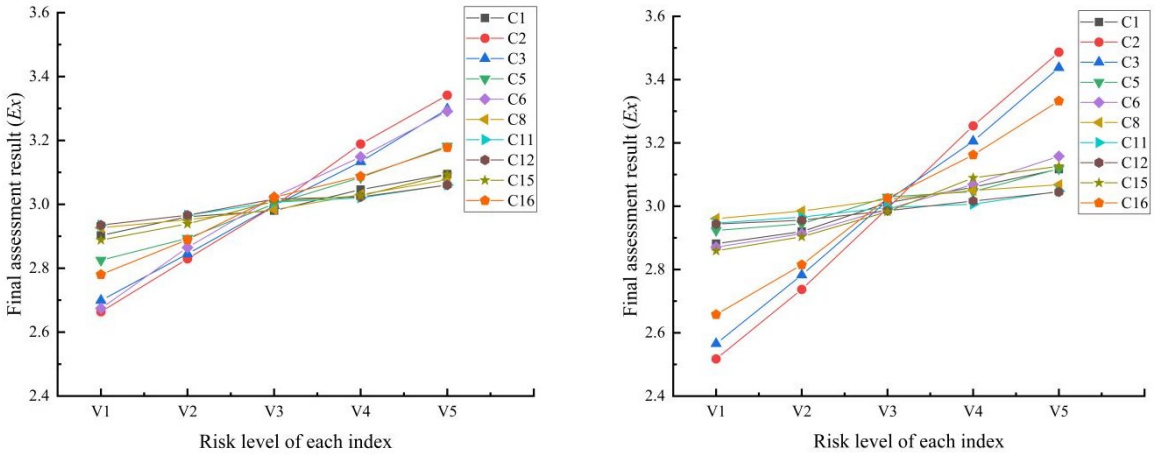
3.14. The other six indexes' sensitivities are very low and are not presented here. The sensitivities of third-level indexes are ranked in Table 3.24. The sensitivities of second-level indexes are calculated using the following equation, and ranked in Table 3.13 and Table 3.23.

$$S(Ci)=R_{V5} - R_{V1} \tag{3.28}$$

where $S(C_i)$ denotes the sensitivity of index C_i, R_{V1} denotes the final result after 1000 simulations when the risk level of the specified index is $V1$, R_{V1} denotes the final result after 1000 simulations when the risk level of the specified index is $V5$.

Table 3.24 Ranking of third-level indexes' sensitivities

	Third-level indexes
Constant weight	C2>C6>C3>C16>C5>C15>C1>C8>C12>C11>C13>C4>C7>C10>C9>C14
Variable weight	C2>C3>C16>C6>C15>C1>C5>C8>C12>C11>C4>C7>C13>C10>C9>C14



(a) Constant weight (b) Variable weight

Figure 3.14 Sensitivity analysis of the third-level indexes

In Figure 3.14, according to the rules of sensitivity calculation, the greater the slope of each dash, the more significant the sensitivity of its corresponding third-level index. Thus, we find that after the optimization of the variable weight function, the sensitivities of indicators C2, C3, C16 are significantly increased compared with the values calculated by the constant weights, indicating that the variable weight function model better highlights the key risk factors in the system.

According to Table 3.13 and Table 3.23, the rank of the weights of second-level indexes for the constant weight method is: B1(corrosion) > B2(third-party damage) > B6(consequences) > B4(safety management) > B3(design and manufacturing) >B5 (operation and management); for the variable weight method, the rank of the weights of second-level indexes is: B1 > B6 > B2 > B4 > B3 >B5. Among them, corrosion, third-party damage, and consequences account for a relatively large proportion of pipeline risk. Because the pipeline is near a subway line, stray currents from the rail transit traction system leads to severe corrosion of the pipeline. In the constant weight ranking, B2 > B6 is because third-party damage index has a greater impact on pipeline system risk than consequence index. However, this ranking is reversed in the variable weight calculation because the consequences index has a smaller risk comment value, implying a higher risk. Therefore, the weight of B6 increased after applying the variable weighting function. According to Table 25, the ranking order of the third-level indexes also changed. C3 and C16 are highlighted, while C6 and C5 are considered to be less important.

The actual pipeline is situated under the side of a busy highway in Beijing, so while there is less risk of third-party damage, dense population and high social sensitivity increase the risk of serious consequences. Therefore, the results obtained using VWT are more consistent with the actual environmental realities surrounding the pipeline. When applying the constant weight method, some comprehensive assessment results may remain relatively safe or acceptable even when some indexes reach the dangerous boundary. However, when applying variable weight theory, if there is an index with a small weight but significant risk, the weight of this index can be increased, thereby reducing the system's risk comment value reasonably and making the assessment result relatively dangerous. This is very helpful for pipeline operators to identify the key factors affecting system safety and can provide guidance for gas pipeline risk management and decision making.

3.4.3 Risk management based on the assessment result

One of the purposes of risk assessment is to rank the relative risks of the pipe sections of a pipeline. The pipeline operators can identify high, medium, and low-risk pipe sections based on the ranking and then determine whether pipe sections require further risk control or risk rehabilitation measures. Thus the mitigation and maintenance activities may be prioritized and will become more cost effective.

We selected the pipeline to which the pipe section studied in this chapter belongs and divided it into 19 pipe sections. Risk assessment and classification are performed for each pipe section, as shown in Table 3.25. Pareto's Law [195] can be used for better statistical purposes. We assign scores to the risk levels. According to Table 3.2, *V1* corresponds to the highest risk level, and so a score of 3 is assigned as the highest risk score. Similarly, *V2* is assigned a score of 2 and *V3* a score of 1. The risks of *V4* and *V5* are acceptable and tolerable, and therefore, both of them are assigned a score of 0.

Table 3.25 Risk assessment results of the pipe sections

Pipe Section No.	Risk Level	Risk Score	Pipe Section No.	Risk Level	Risk Score
0010101	*V2*	2	0010111	*V4*	0
0010102	*V1*	3	0010112	*V5*	0
0010103	*V4*	0	0010113	*V4*	0
0010104	*V3*	1	0010114	*V3*	1
0010105	*V2*	2	0010115	*V5*	0
0010106	*V5*	0	0010116	*V1*	3
0010107	*V4*	0	0010117	*V4*	0
0010108	*V5*	0	0010118	*V4*	0
0010109	*V2*	2	0010119	*V5*	0
0010110	*V4*	0	Sum of risk score		14

As shown in Table 3.25, there are five pipe sections (No. 0010101, No. 0010102, No. 0010105, No. 0010109, No. 0010116) with higher risk, which is 26.32% of all pipe sections, accounting for 85.71% of the total risk score. In other words, most of the risks in this pipeline concentrate on the five pipe sections. Suppose the pipeline operators pay attention to 26.32% of pipe sections. In such a case, 85.71% risk can effectively be controlled, which would significantly optimize the risk mitigation resources and increase the economic efficiency of the pipeline. In the meantime, the ranking of pipe

sections' risk levels can be considered a prioritization when performing risk rehabilitation on this pipeline. Higher risk pipe sections should be prioritized for targeted control measures, depending on the key risk factors identified during the evaluation process. When two pipe sections have the same *Ex* value, the value of *En* should be evaluated. The larger the value of *En*, the greater the uncertainty and the higher the risk.

For some high-risk pipe sections where risk mitigation is very difficult, it is impossible to carry out risk mitigation immediately after completing the risk assessment. Therefore, pipeline operators need to make emergency plans based on the actual situation of the pipeline. This will minimize the damage even if the worst outcome occurs.

3.5 Conclusions

This research aims to develop a new risk assessment method for gas pipelines. This chapter adopts a multi-factor coupled risk analysis method that comprehensively considers 16 factors affecting the pipeline's safe operation. These factors are divided into a three-level risk assessment index system, ensuring that the assessment results are more reliable through level-by-level calculation.

This chapter proposes a novel cloud-variable weight theory and applies cloud theory throughout the entire risk evaluation process, effectively expressing the uncertainty in the process. The CM is used to express constant weights, variable weights, and expert comments, eliminating the error in the existing methods for single numerical modeling of fuzzy concepts. At the same time, the constant weight calculation method based on cloud-AHP and virtual floating cloud computation can express uncertainty while increasing the data aggregation, thereby making the assessment results more reasonable.

In this chapter, cloud theory and variable weight theory are well integrated and improve each other. As a result, the weights of the second-level index are optimized, thus highlighting the negative effects of the index with higher risk, reducing the wrong

judgment caused by the constant weight method, and providing a more scientific and reasonable method for pipeline risk assessment.

After calculating the industrial case, we found that 26.32% of all pipe sections account for 85.71% of the pipeline's total risk. Therefore, the proposed method can help pipeline operators make optimal risk decisions with limited resources. Furthermore, for the pipe sections specifically analyzed in Section 3, among all the indexes, the corrosion index accounts for the largest proportion of pipe section risk, which should be focused on during the risk control process. Thus the pipeline operators can determine the prioritization when performing risk rehabilitation and the weak links to implement targeted risk management strategies.

The proposed method is a promising approach to risk analysis of the buried gas pipeline. However, some limitations are summarized as follows: (1) the established risk assessment index system is only applicable to buried gas pipelines, but not to other types such as submarine pipelines or long-distance oil pipelines; (2) the proposed method is a static approach that only analyzes risk based on the current state of the pipeline and cannot dynamically track the development of risk. Therefore, future research should focus on these problems to better ensure safe and smooth operation of pipelines.

Chapter 4: Discovery of potential risks for the gas transmission station using monitoring data and the OOBN method

4.1 Introduction

For the urban consumers along the long-distance pipeline, gas transmission stations (GTS) are essential elements that filter and separate impurities, regulate flow and pressure, and distribute gas [196, 197]. Due to the complexity of the GTS process and the high flammability and explosibility of the transported media, some risk factors, such as corrosion, misoperation, etc., may lead to equipment failure or even accidents [198]. For example, on January 20, 2006, an explosive accident at the Fuga gas station in China killed 10 people and injured 50 [199]. On June 23, 2012, ignition and fire occurred in British Columbia, Canada. The number of injuries reported was 2 [200]. Between 1994 and 2013, there were 745 significant gas distribution accidents in the United States, resulting in 278 deaths and 1059 injuries, as well as $110,658,083 in property damage [201].

The severity of an accident's consequences emphasizes the need for risk management for GTS. In the case of an accident, the gas transmission facilities are located in urban or residential areas, making it harder to evacuate people and take emergency steps to safeguard property [202, 203]. Identifying accident precursors and preventing accidents before they happen is thus the least costly and most effective risk management strategy for GTS operators. In existing engineering practice, the most common method for early warning is combustible gas detectors. However, this method can only provide early warning for indoor leaks while significantly less capable of detecting outdoor leaks and non-leakage faults such as blocking. Therefore, it is necessary to further improve the ability of the GTS system to discover potential risks

in real-time.

The combination of accident precursor analysis and quantitative risk analysis has been investigated in [204], in which accessible information from accident precursors is utilized as input to the QRA technique to determine potential consequences. The definition of an accident precursor is “an abnormality that suggests the possibility for more severe effects that may occur in the future, due to factors observable from its existence today” [205]. Identifying system abnormalities, therefore, aids in the detection of accident precursors. Furthermore, equipment failure leading to pressure anomalies has been demonstrated in [206-208]. GTS’s supervisory control and data acquisition (SCADA) system can quickly acquire pressure and flow data, which is the monitoring data we used in this study. The operators can then decipher the pattern of data changes to identify accident precursors and achieve real-time risk management through modeling. Currently, no real-time potential risk analysis methods for GTS systems have been proposed to the authors’ knowledge. Moreover, since the GTS system is complicated, with some duplicated structures, limited data sources, and some missing records in practice, existing real-time risk analysis methods [209-211] in other fields are not applicable to the GTS system.

Table 4.1 summarizes the existing risk studies for GTS and other energy transmission and distribution facilities similar to GTS. We can conclude from these works that (1) BN is the most commonly used risk analysis approach, and most studies [49, 199, 212-215] directly map the established BTA, FTA, or ETA models to BN, including structure mapping and CPT parameter mapping. The model’s complexity and the massive number of nodes make manipulation and analysis difficult. The OOBN model presented in [126], on the other hand, has significant advantages in terms of flexibility and adaptability. Regarding parameter mapping, just the absolute logic of “and” and “or” gates is considered, disregarding any uncertainty. In addition, the multilevel BN method proposed in [25] is based on the perspective of multi-flow intersection, and the Stochastic Petri nets method is used in [26] to analyze GTS’s leakage emergency disposal process. Both propose innovative perspectives in GTS’s

risk research. (2) Since risk changes over time, some studies have employed BN's updating mechanism to conduct a dynamic risk analysis [49, 126, 213, 214]. However, because the data required for these updates are difficult to obtain in a short period, only periodic updates of risks, rather than real-time updates to sensitively identify accident precursors, can be undertaken. (3) In terms of posterior analysis of risk factors, most existing methods conduct backward analysis assuming that an accident has occurred. Therefore, there must be an available data source to support this assumption. Expert knowledge acquisition used in [49, 126, 199] and numerical simulations in [216] generally need a significant amount of time and human resources. In contrast, data from databases and literature [199, 213, 214] are generic and not very targeted and precise due to the complexity of the process. And [217] classifies monitoring data into intervals, losing some of its accuracies.

Table 4.1 Some relevant studies

Literature	Risk analysis method	Data acquisition method
[199]	Bow-tie analysis (BTA), Bayesian network (BN)	General database and expert knowledge
[212]	Fault tree analysis (FTA), BN	Calculate failure probability by mean time to failure (MTTF) in historical data
[49]	Failure mode and effect analysis (FMEA), BTA, BN	Expert knowledge
[213]	FTA, event tree analysis (ETA), BN	General database
[214]	FTA, event sequence diagram, BN	Reliability database, literature, and cumulative abnormal event data
[215]	BN, Layer of Protection Analysis (LOPA)	Plant-specific failure data
[126]	OOBN	Expert knowledge
[216]	Grid-based method, BN	Numerical simulation using DNV PHAST
[217]	Fuzzy expert system	Translate monitoring and detection data into linguistic terms and fuzzy functions
[218]	Multi-flow intersecting theory, multilevel BN	Expert knowledge
[219]	Stochastic Petri nets	Expert knowledge

For the GTS systems' actual situation and the limitations of the methods mentioned above, a method for potential risk discovery in GTS systems is proposed in this research. In comparison to previous research, the contributions are summarized as

follows. (1) A structure mapping method based on FMEA is proposed. (2) An OOBN framework is developed based on the overall system's process flow, making the model more simplified and flexible. For example, compared to the BN framework, this will reduce the number of nodes by over 50% and the number of conditional probability tables (CPTs) by over 60% in the examples investigated in the paper. The operators can call up the built model for duplicate structures in the system instead of re-modeling it. (3) Both leaky noisy-or gate (LNG) and expectation maximum (EM) algorithm are incorporated into BN's parameter learning, efficiently dealing with imperfect historical records and limited sample data while reflecting the uncertainty of risk factors. (4) An accident precursor identification approach based on PAA-CUSUM is presented to identify possible vulnerabilities in the system in real-time by finding abnormalities using monitoring data. This allows backward analysis when anomalies are detected, which is more consistent with actual conditions than assumption-based analyses.

The rest of this work is structured as follows. Section 2 introduces some basic concepts and the proposed method; Section 3 describes modeling the GTS system; Section 4 selects some cases to illustrate and validate the proposed method; Section 5 provides a more in-depth discussion of the proposed method; Section 6 presents the conclusion.

4.2 Methodology

4.2.1 Bayesian network

Bayesian network (BN) is the result of combining graph theory and probability theory, and their benefits in dealing with uncertainty have landed them a spot in a variety of risk analysis methodologies [220, 221]. BN is defined as G = ((*N*, *A*), *P*), where (*N*, *A*) represents a graph, *N* is a node, *A* is a directed arc, and *P* represents the conditional probability table (CPT) associated with each node. The joint probability is represented by the interrelationship of the parent nodes and the conditional probability, as shown in Eq. (4.1) [222].

$$P(X_1, \cdots, X_n) = \prod_{i=1}^{n} P(X_i | Pa(X_i)) \quad (4.1)$$

where $X=\{X_1,\ldots X_i,\ldots, X_n\}$ represents the variable set, $Pa(X_i)$ is the parent node of X_i, if $Pa(X_i)$ is an empty set, then X_i is the root node, $P(X_i \mid Pa(X_i))=P(X_i)$ represents its prior probability.

4.2.2 Object-oriented Bayesian network

There is a large variety of process flow in complex systems, such as GTS systems, since they involve various pipes, valves, and other equipment and devices. On the other hand, some subsystems have close or even identical structures. As a result, integrating object-oriented concepts into Bayesian networks can minimize the complexity of traditional Bayesian network modeling while also increasing the reusability of model pieces [223-225]. The idea of classes is introduced in OOBN. Each object is an instantiated representation of a class, which defines a group of objects with the same structure and behavior [226]. Objects 1 and 5 are instantiated based on class A, 3 and 6 are instantiated based on class B, and object 2 is instantiated based on class C, as indicated in Figure 4.1. In conclusion, the OOBN model for complex systems comprises BN pieces representing each subsystem. These BN pieces communicate with the outside environment through input and output nodes. The input data contain pre-processed historical data, sensor data, and expert experience and knowledge. After an algorithmic model analysis, the system's risk and functional condition can be output. Internal nodes only have parents and children in their related BN fragments, not in other BN fragments, due to the encapsulated nature of the class.

OOBN provides the following benefits over traditional BNs [227]. After the overall network has been validated, OOBN provides a top-down modeling approach. Each Bayesian network fragment may be enhanced step by step. OOBN builds complicated models from small, easy-to-understand model fragments, making expert knowledge acquisition and communication between modelers and experts easier. The OOBN method altogether includes encapsulation and hierarchical properties, and it has a faster convergence time and efficiency than the traditional Bayesian network

inference technique [228].

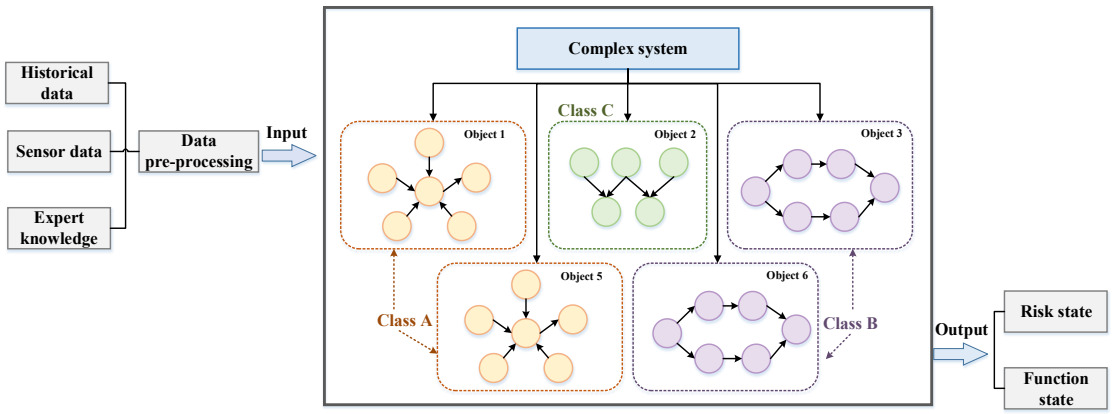


Figure 4.1 OOBN-based modeling methodology

4.2.3 Proposed method

4.2.3.1 Overview of the proposed method

The specific steps of the proposed method are shown in Figure 4.2. Firstly, the complex system is divided into subsystems defined as objects. Then BN structure of the objects is developed, and the parameters of the CPTs are determined to form an OOBN model of the whole system. Based on the PAA-CUSUM algorithm, the change degrees of indicator readings are calculated, and the objects with potential risks are identified and ranked according to the change degree magnitude. Finally, the components' accident precursors and potential risk ranking are identified according to the calculation of the posterior and prior probabilities.

In this study, OOBN can model complex systems by dividing them into subsystems based on the process flow, and repetitive systems only need to be modeled once in the study and then called up when needed. Compared with existing methods in Table 4.1 Some relevant studies, the proposed approach is more flexible and straightforward. It considers the nature of various subsystems in complex systems, so that specific high-risk subsystems and components can be accurately located. Moreover, the trend of the data after PAA-CUSUM processing is used as input to the model, rather than just a classification of sensor data by numerical magnitude, which can result in a

more effective discovery of the potential risks. In addition, compared to existing methods such as principal component analysis (PCA) [229], the data processing results of PAA-CUSUM are interpretable and, therefore, more easily accepted and understood by users.

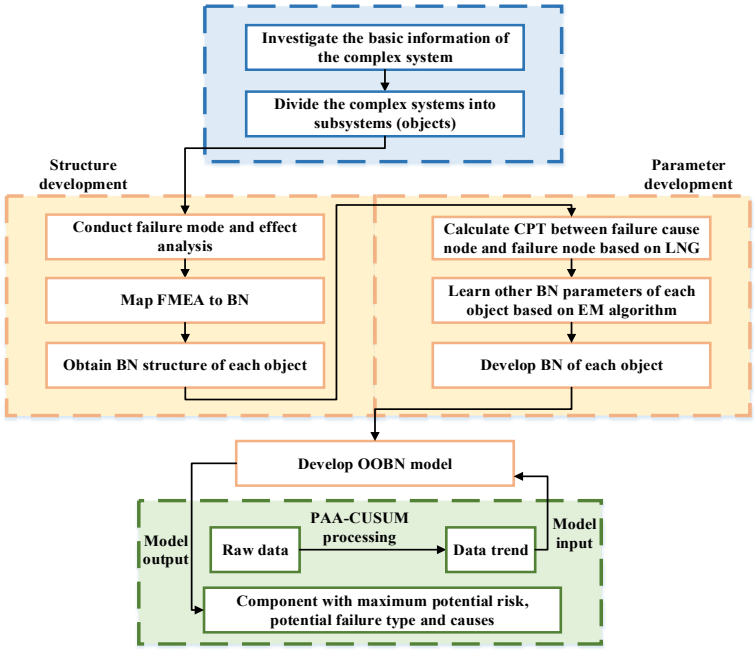


Figure 4.2 Schematic diagram of the proposed method

4.2.3.2 BN structure development

Expert knowledge, machine learning techniques, or a combination of the two approaches are the most common ways to structural modeling of BNs [230]. Large systems typically have complicated structures and various failure modes, and structural modeling using machine learning may generate conclusions that violate engineering conventions. FMEA procedures, on the other hand, help detect potential failure modes and effects in complex systems [231-233]. Compared to existing mapping methods based on BTA [199] and FTA [213], FMEA-based mapping is more user-friendly for GTS users because it can reduce the number of nodes while serving the same function of cause-and-effect analysis. As a result, using FMEA mapping, this chapter proposes an expert knowledge-based technique for determining the structure of subsystem Bayesian networks. The mapping approach combines FMEA's layer-by-layer

characteristic with Bayesian networks' cause-and-effect framework, allowing complex system administrators to examine problems more systematically and logically.

The root node of BN corresponds to the cause of failure determined by FMEA. The first-level intermediate node shows whether the component's operational state is normal or not, as illustrated in Figure 4.3.

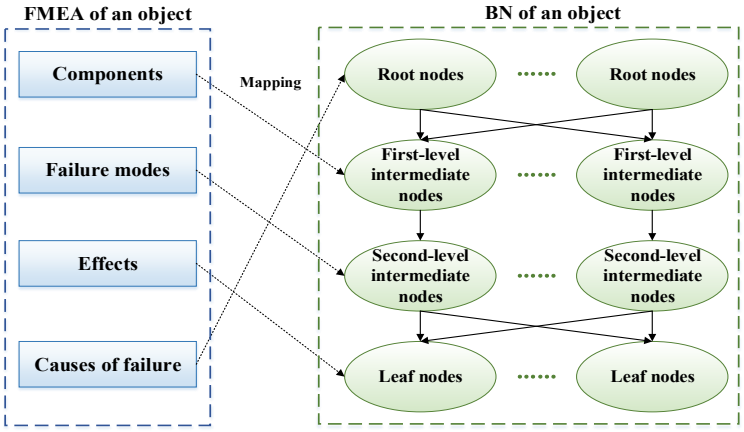


Figure 4.3 Mapping rules from FMEA to BN

As shown in Figure 4.3, the failure mode identification of FMEA can determine the failure type of each component represented by the second-level intermediate node, and the leaf nodes can reflect the impact of different failure types.

4.2.3.3 BN parameter determination

In order to better solve the problems of missing records, small sample data size, and uncertainty of risk factors in engineering practice, this study combines two methods, EM algorithm and LNG, together to determine the parameters of BN. The Noisy-or gate (NG) is a common paradigm for describing the relationship between causes and their associated co-influence [234]. The essential premise of this gate is that any causative factor, even in the absence of other causes, can independently influence the common result. 2^N separate conditional probability parameters must be collected for a child node Y_1 with N parents (X_1, …, X_N), yet obtaining all of this information in practice is difficult. In contrast, the Noisy-or model only requires $2N$ conditional probability parameters, considerably increasing the efficiency of acquiring probability parameters from historical data and expert knowledge. In this case, it is commonly assumed that each

parent node's effect on Y_1 is independent, i.e., the nodes in the BN are regarded as Noisy-Or nodes, and the value of Y_1 is determined by the contribution of each parent node to it via the "or" operation in Eq.(4.2) [235].

$$Y_1 = \varepsilon_1 \text{ V } \varepsilon_2 \text{ V } \dots \text{ V } \varepsilon_N \tag{4.2}$$

where Y_1 denotes the child node, ε_i denotes the contribution of parent node X_i to Y_1 .

$$P(\varepsilon_i = 1|X_i = 1) = P(Y_1 = 1|X_1 = 0, \cdots, X_i = 1, \cdots, X_N = 0) \tag{4.3}$$

Then the conditional probability of obtaining Y_1 from causes $X_1, X_2, ..., X_N$ is shown in Eq.(4.4).

$$P(Y_1|X_P) = 1 - \prod_{X_i \in X_N}(1 - p_i) \tag{4.4}$$

where X_P denotes the parent node set, p_i denotes the joint probability of X_i.

The joint probability p_i can be calculated by Eq. (4.5), and the specific derivation process is shown in [236].

$$p_i = \frac{P(Y_1|X_i) - P(Y_1|\overline{X_i})}{1 - P(Y_1|\overline{X_i})} \tag{4.5}$$

Some other unknown factors on Y_1 may be gathered into a factor X_L, which is also added to the network as a parent node to construct a Leaky Noisy-or gate (LNG), thereby boosting the expression of uncertainty in the model [237]. L denotes the subscript of the set of unknown factors, which follows a Gaussian probability distribution with a confidence level of 99% [238, 239]. The conditional probability that the child node is true is expressed as:

$$P(Y_1 = 1) = 1 - \prod_{X_i \in X_N}(1 - p_i)(1 - p_L) \tag{4.6}$$

where p_L denotes the leak probability, $p_L = P(Y_1 = 1|X_1 = 0, \cdots, X_i = 0, \cdots, X_N = 0, X_L = 1)$.

The Expectation-Maximum (EM) algorithm is an iterative approach to solving the maximum value issue [240]. Each iteration of the method is broken into two parts, E and M. The implied data and model distribution parameters are updated via several iterations until convergence is achieved, after which the model parameters are obtained [241]. The specific steps are as follows.

EM algorithm

Inputs: Predefined sample data set $x=(x^{(1)}, x^{(2)},..., x^{(m)})$, unobserved implied data set $z=(z^{(1)}, z^{(2)},..., z^{(m)})$, maximum number of iterations j.

Outputs: Model parameters θ.

Step 1: Randomly initialize the model parameter θ^0.

Step 2: Compute the conditional probability expectation of the joint distribution by Eq.(4.7) and Eq.(4.8).

Step 3: Maximize $L(\theta,\theta_j)$ to obtain θ_{j+1}: $\theta^{j+1} = \underset{\theta}{\operatorname{argmax}} L(\theta, \theta^j)$

Step 4: If θ_{j+1} has converged, the algorithm ends. Otherwise, continue back to step 2 for iteration.

$$Q_i\left(z^{(i)}\right) = P(z^{(i)}|x^{(i)}, \theta^j) \tag{4.7}$$

$$L\left(\theta, \theta^j\right) = \sum_{i=1}^{m} \sum_{z^{(t)}} Q_i(z^{(i)}) logP(x^{(i)}, z^{(i)}; \theta) \tag{4.8}$$

where $L(\theta,\theta_j)$ denotes the likelihood function, $Q_i\left(z^{(i)}\right)$ represents distribution of the implied variable z for sample i.

Because the data in this study originate from accident reports, there will be some imperfect records and limited sample size. Through a quicker convergence rate, the EM algorithm can adequately estimate the probability value between nodes under the premise of a limited number of samples, which can compensate for this shortcoming [242]. As a result, the EM method is chosen as the parameter learning algorithm.

4.2.3.4 Accident precursor identification

The piecewise aggregate approximation (PAA) algorithm for time series data dimensionality reduction has proven effective. PAA operation is straightforward and more appropriate for engineering applications than other dimensionality reduction approaches [243]. The main principle behind it is as follows.

Let the length of a time series $X=\{x_1, x_2,..., x_n\}$ be n, such as pressure or temperature data. Representing it as a vector $\bar{X}$ of length N, the i-th element is defined as:

$$\bar{X}_i = \frac{N}{n}\sum_{j=\frac{n}{N}(i-1)+1}^{\frac{n}{N}i} x_j \tag{4.9}$$

The PAA method can improve the efficiency of time series anomaly detection and reduce the effect of noise by reducing the time series from *n*-dimensional space to *N*-dimensional space.

The CUSUM algorithm is an effective method for keeping track of industrial irregularities. For example, it can identify possible faults in wind turbine main bearings and other circumstances that might produce changes in the signal [244]. The cumulative sum is the total amount of deviations, including data from all prior samples. They're great at detecting slight changes in the variable [245]. The CUSUM method is used to analyze observations that have been collected over time. These observations might be physical measurements, counts, or ratios, and they can be grouped (e.g., in the form of manufacturing batches) or individual observations [246].

Based on the characteristics of the signal sequence in the statistics, the CUSUM algorithm assesses if the change points exist in the monitoring process. The system generates a change point when the cumulative sum is considerably greater or lower than the usual operating condition [247]. A signal sequence's cumulative sum is a variable that fluctuates randomly around its initial value until it changes abruptly. If a positive shift happens at a change point, CU_i continues to increase, indicating positive accumulation. Vice versa. It can be determined that a signal mutation has been formed when it has accumulated to a certain level (i.e., more than a pre-set threshold).

The followings are the steps in the approach for change point identification based on the PAA-CUSUM algorithm proposed in this chapter, shown in Figure 4.4.

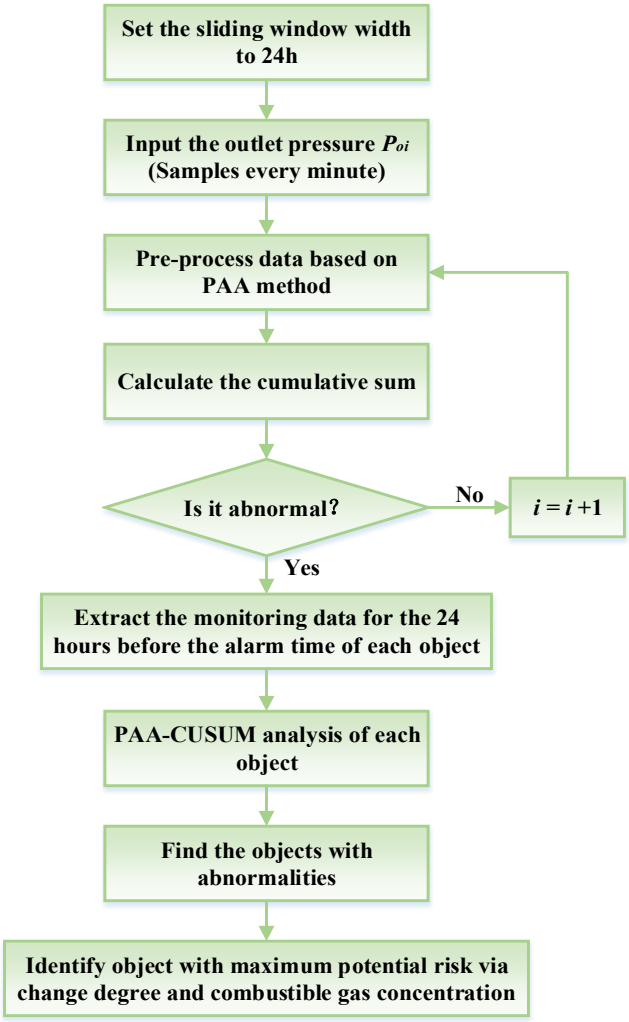


Figure 4.4 Fault location process

Step 1: Implement the sliding window approach suggested in [248] and set the length as $w = 24$ hours. After the accident precursor has been located, the PAA-CUSUM algorithm can be reset for a new cycle.

Step 2: Use the PAA algorithm to pre-process the pressure data at the system's outlet. Calculate the lower and upper cumulative sums and boundary values using Eq. (4.10)-(4.13) [249].

$$CL_i = \min\left(0, CL_{i-1} + x_i - \left(T - k\frac{\sigma}{\sqrt{m}}\right)\right) \tag{4.10}$$

$$CU_i = \max\left(0, CU_{i-1} + x_i - \left(T + k\frac{\sigma}{\sqrt{m}}\right)\right) \tag{4.11}$$

$$UCL = \frac{Nh}{n} \times \frac{\sigma}{\sqrt{m}} \tag{4.12}$$

$$LCL = -\frac{Nh}{n} \times \frac{\sigma}{\sqrt{m}} \tag{4.13}$$

where i denotes the i-th element after pre-processed by PAA algorithm, CL_i denotes the lower cumulative sum, CU_i denotes the upper cumulative sum, UCL and LCL are the upper and lower boundary values, respectively. T is the target value, k is the size of the shift to be detected, h is the standardized decision interval, σ is the variable in control, m is number of subgroups.

Step 3: Determine whether there is an abnormality. In order to reduce the rate of misjudgment and omission of the system, the criteria for abnormality are (1) when there is a downward or upward trend (some consecutive points increase or decrease), and (2) when the alert threshold value is set according to ISO standard (ISO 7870-4) is exceeded [250]. In this chapter, to be more consistent with the actual situation, we choose m=1, k=0.5, and h=3.5. The system is judged to be abnormal if both are satisfied. If there is an abnormality, proceed to the fourth step. If there is no abnormality, return to step 2.

Step 4: Retrieve the monitoring data of each object for 24 hours before the alarm time. Sum up the CL_i and UL_i and plot the PAA-CUSUM chart depending on which one is more extreme.

Step 5: Find the object with anomaly and calculate the change degree of its CUSUM chart based on the Eq. (4.14)-(4.15) adapted from [251].

$$P_u(\{x_j\}|\{x_i\}) = \prod_{j=1}^{l} P(X \geq X_j|\{x_i\}) \tag{4.14}$$

$$P_d(\{x_j\}|\{x_i\}) = \prod_{j=1}^{l} P(X \leq X_j|\{x_i\}) \tag{4.15}$$

where $\{x_i\}$ is the data set before the change point, $\{x_j\}$ is the data set after the change point, which both follow the Gaussian distribution. P_u denotes the probability of an upward trend in change after the observation of $\{x_i\}$. Similarly, P_d, is the probability of a downward trend. L denotes the number of elements in the set $\{x_j\}$. To calculate upward change u and downward change d more intuitively, the formulas are transformed into Eq. (4.16)-(4.17).

$$u = -\frac{1}{l}\sum_{j=1}^{l} logP(X \geq x_j\,|\{x_i\}) \tag{4.16}$$

$$d = -\frac{1}{l}\sum_{j=1}^{l} logP(X \leq x_j \,|\{x_i\}) \tag{4.17}$$

Step 6: Find the object with the largest change degree and identify the accident precursor. If the chart trend is continuously decreasing, it means that the system may have a leak, then in engineering practice, the handheld combustible gas concentration detector can be used to assist in diagnosis and improve accuracy.

4.2.3.5 Potential risk analysis

The upward trend or downward trend identified in section 2.3.4 is used as the input of each object's BN fragment, which is named as "lower," "normal," or "higher." The difference Δp between the posterior and prior probability of BN is used to measure potential risk and provide risk ranking [252, 253]. For GTS operators, this ranking can be used to determine the order of risk remediation and improve the efficiency of daily maintenance.

4.3 Case study

4.3.1 Objects identification

An industrial case study exemplifies the proposed methodology's various steps. A gas transmission station (GTS) in Shaanxi, China, is used as example. The basic process flow diagram of the station is shown in Figure 4.5. According to the experts' opinion and the GTS's key equipment functions mentioned in [4, 22], this GTS is divided into five categories of units: entrance unit, filtration and separation unit (FSU), pressure regulating unit (PRU), flow metering unit (FMU), and export unit. Based on the overall structure of the gas transmission station, we simplify the process flow into a reliability block diagram (RBD), as shown in Figure 4.6. For example, there are two objects with identical structures in the FSU. Then, in RBD, they are simplified as object 2 and object 3. Similarly, the two objects with identical structures in PRU are simplified as object 4 and object 5; FMU1 in Figure 4.5 is simplified into object 7 and object 8; FMU2 in

Figure 4.5 is simplified into object 6. It should be mentioned that this chapter only studies the situation when the gas transmission station is in regular operation, so the venting pipeline and valves are not reflected in the diagram. Due to the limitation of the diagram, we cannot put all the gas lines and drainage lines into the dashed boxes of the corresponding units in the diagram. Still, they are taken into account in the actual analysis.

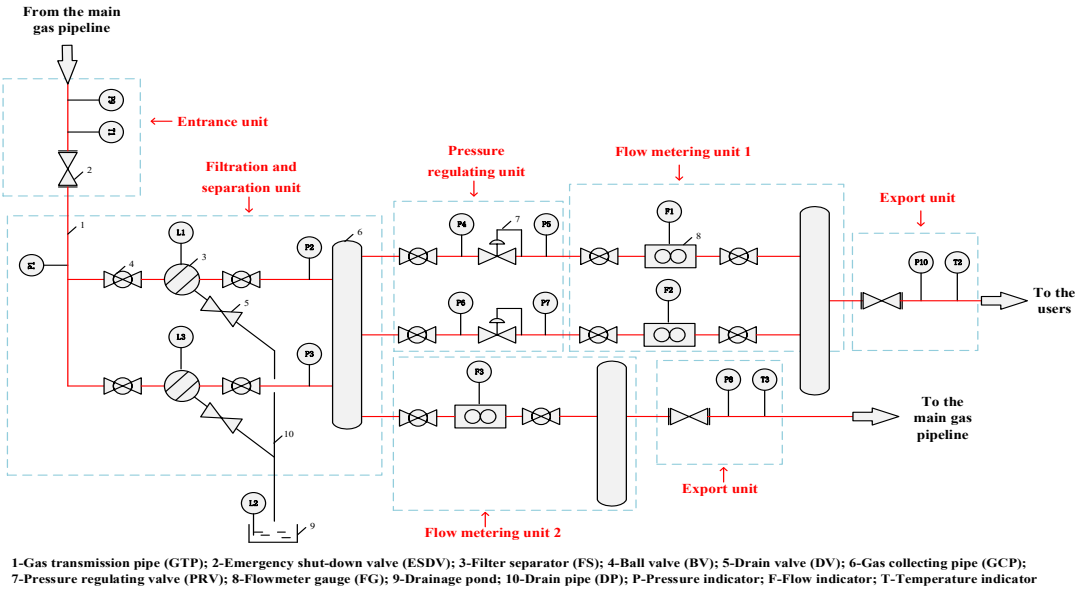


Figure 4.5 Process flow diagram

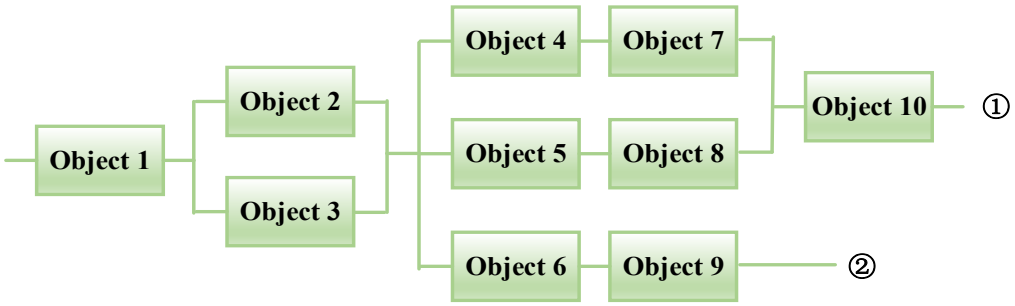


Figure 4.6 Reliability block diagram

4.3.2 BN structure modeling of each object

Take the filtration and separation unit as an example. According to the procedure of FMEA, firstly, we draw a functional block diagram, as shown in Figure 4.7. Then we conduct FMEA based on the identified functions.

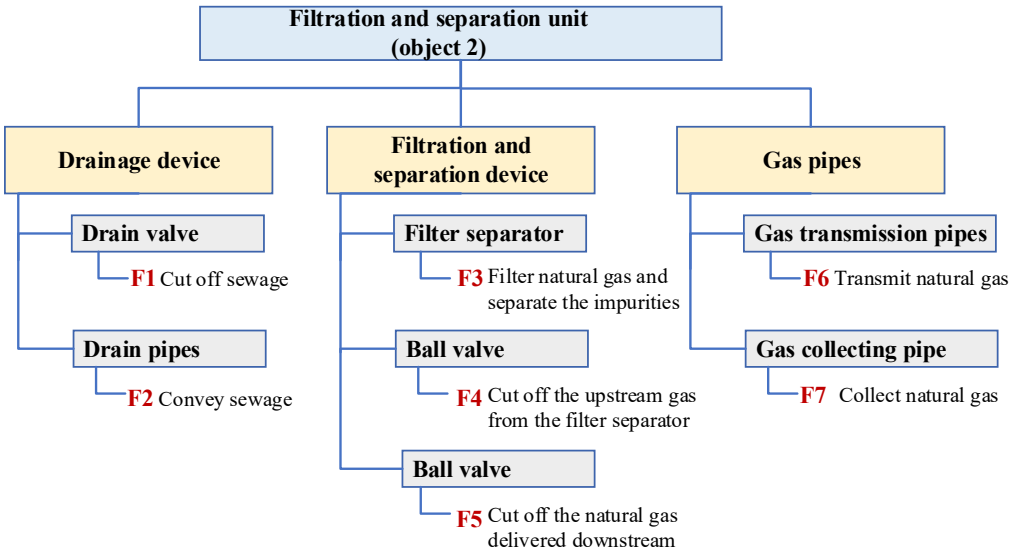


Figure 4.7 Functional block diagram

Corrosion can cause loss of metal and thus cause leakage. Also, corrosion can lead to valve rust, causing valve components to jam. The valves and pipes can easily form freeze blockage in a low-temperature environment. And in normal environments, foundation settlement can also affect the deformation or failure of components. Management factors can also lead to failures, such as insufficient valve lubrication, design, and construction defects [199]. According to Table 4.2, we map FMEA to BN, as shown in Figure 4.8.

Table 4.2 FMEA of object 2

Function	Component	Failure mode	Effect	Causes of failure
F1	Drain valve	Internal leakage	The valve cannot be closed tightly, L1 is lower than normal, and L2 is higher than normal.	Valve flap corrosion; improper assembly of components
		External leakage	Drainage leaks, L1 and L2, are lower than normal.	Valve stem corrosion; deformation of the valve and pipe connection due to foundation settlement; low temperature causes components to freeze and crack.
		Blocking	The drainage system cannot operate properly, L1 is higher than normal, and L2 is lower than normal.	Lubrication failure; actuator failure; component rusting; low temperature
F2	Drain pipes	Leakage	Drainage leaks, L1 and L2, are	Corrosion; foundation

			lower than normal	settlement; external damage
		Blocking	The drainage system cannot operate properly, L1 is higher than normal, and L2 is lower than normal.	Low temperature
F3	Filter separator	Cartridge blocking	The filtration system cannot operate properly; L1 is higher than normal, P1 is higher than normal, and P2 is lower than normal.	Cartridge aging; excessive impurities in natural gas; unqualified cartridge; low temperature
		Leakage	Media leaks, L1, P1, and P2 are lower than normal.	Cylinder with an unqualified seal; corrosion; foundation settlement
F4	Left ball valve	Internal leakage	The valve cannot be closed tightly, P1 is lower than normal, and P2 is higher than normal.	Valve flap corrosion; improper assembly of components
		External leakage	Gas leaks, P1 and P2, are lower than normal.	Valve stem corrosion; foundation settlement; low temperature
		Blocking	The system cannot operate properly, P1 is higher than normal, and P2 is lower than normal.	Lubrication failure; actuator failure; component rusting
F5	Right ball valve	Same as F4		
F6	Gas transmission pipe	Leakage	Gas leaks, P1 and P2, are lower than normal	Corrosion; foundation settlement; external damage; low temperature
F7	Gas collecting pipe	Leakage	Gas leaks, P1 and P2, are lower than normal	Corrosion; foundation settlement; external damage; low temperature

The causes of failure in Table 4.2 are summarized and divided into four categories: corrosion control level, foundation settlement, environmental temperature, and management level. The “as low as reasonably practicable” (ALARP) concept divides the risk into three regions: unacceptable, tolerable, and acceptable. Therefore, for the risk of foundation settlement, we simplify the degree into two levels: acceptable and unacceptable; for corrosion control, environmental temperature, and management, “normal” and “low” would be more suitable.

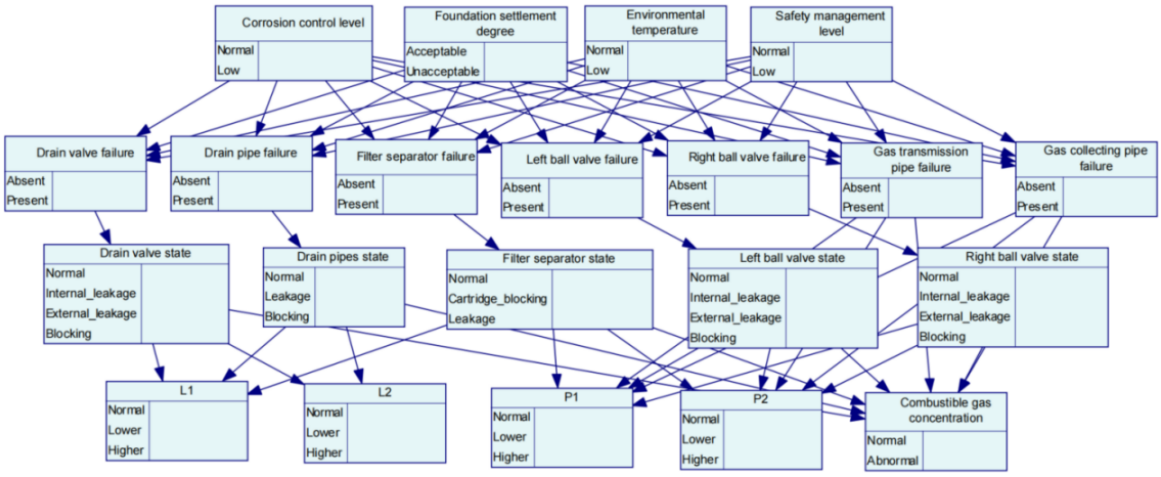


Figure 4.8 BN structure of object 2

4.3.3 BN parameter calculation

LNG is used to analyze the relationship between the failure cause node (first layer of BN) and the failure node (second layer of BN). The initial CPTs are obtained through expert knowledge, as shown in Table 4.3, and the detailed calculation steps are shown in the Appendix. The CPTs using LNG calculation in Eq. (4.6) are shown in Table 4.4. Other nodes are calculated similarly and therefore omitted. The CPTs between failure node and state node are shown in Table 4.5.

Table 4.3 The initial CPTs

Parent	Corrosion control level	Foundation settlement degree	Environmental temperature	Safety management level
State	Normal	Acceptable	Normal	Normal
Absent	0.7703	0.2098	0.5425	0.6184
Present	0.2297	0.7902	0.4575	0.3816

In Table 4.4, there is still a 1% probability that the DV failure state is absent when the risk factors occur. Because in this chapter, we have only used the main risk factors for analysis, but there are still other unknown factors that lead to DV failure, which is called uncertainty. Using Leaky Noisy-or gate, we gather these unknowns into a factor X_L. The known risk factors occur, but X_L does not, resulting in a 1% probability of DV failure to be absent.

Table 4.4 CPTs of DV failure

Corrosion control level	Foundation settlement degree	Environmental temperature	Safety management level	DV failure	
				Absent	Present
Normal	Acceptable	Normal	Normal	0.9686	0.0314
Normal	Acceptable	Normal	Low	0.9177	0.0823
Normal	Acceptable	Low	Normal	0.9314	0.0686
Normal	Acceptable	Low	Low	0.8203	0.1797
Normal	Unacceptable	Normal	Normal	0.9603	0.0397
Normal	Unacceptable	Normal	Low	0.8960	0.1040
Normal	Unacceptable	Low	Normal	0.9132	0.0868
Normal	Unacceptable	Low	Low	0.7726	0.2274
Low	Acceptable	Normal	Normal	0.8634	0.1366
Low	Acceptable	Normal	Low	0.6421	0.3579
Low	Acceptable	Low	Normal	0.7015	0.2985
Low	Acceptable	Low	Low	0.2177	0.7823
Low	Unacceptable	Normal	Normal	0.8272	0.1728
Low	Unacceptable	Normal	Low	0.5471	0.4529
Low	Unacceptable	Low	Normal	0.6222	0.3778
Low	Unacceptable	Low	Low	0.0100	0.9900

As shown in Table 4.5, the EM algorithm calculates the CPTs between other nodes. The original data is taken mainly from the historical failure records of the GTS. In the EM algorithm, the M step: maximum likelihood estimation may be used to overcome the problem of incomplete historical data records.

Table 4.5 CPTs between failure node and state node

Node	State	Present				
		DV failure	DP failure	FS failure	Left BV failure	Right BV failure
DV state	Normal	0	-	-	-	-
	Internal leakage	0.2500	-	-	-	-
	External leakage	0.5033	-	-	-	-
	Blocking	0.2467	-	-	-	-
DP state	Normal	-	0	-	-	-
	Leakage	-	0.5842	-	-	-
	Blocking	-	0.4158	-	-	-

FS state	Normal	-	-	0	-	-
	Cartridge blocking	-	-	0.2801	-	-
	Leakage	-	-	0.7199	-	-
Left BV state	Normal	-	-	-	0	-
	Internal leakage	-	-	-	0.2957	-
	External leakage	-	-	-	0.5524	-
	Blocking	-	-	-	0.1519	-
Right BV state	Normal	-	-	-	-	0
	Internal leakage	-	-	-	-	0.1742
	External leakage	-	-	-	-	0.6105
	Blocking	-	-	-	-	0.2153

4.3.4 OOBN modeling of the GTS system

Similarly, as shown in Figure 4.9 and Figure 4.10, we model the pressure regulating unit and the flow metering unit in the same method as in Section 3.2 and 3.3.

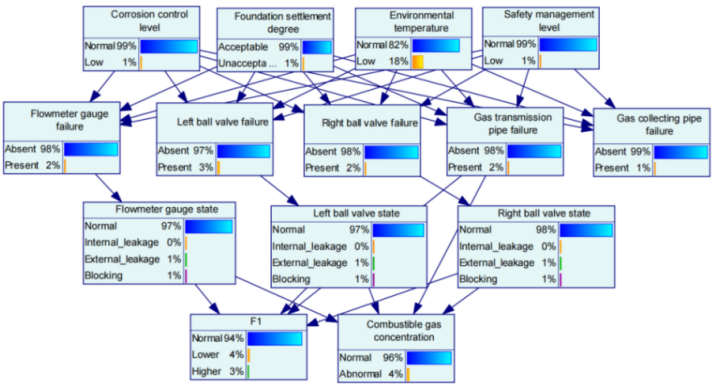


Figure 4.9 BN of object 7

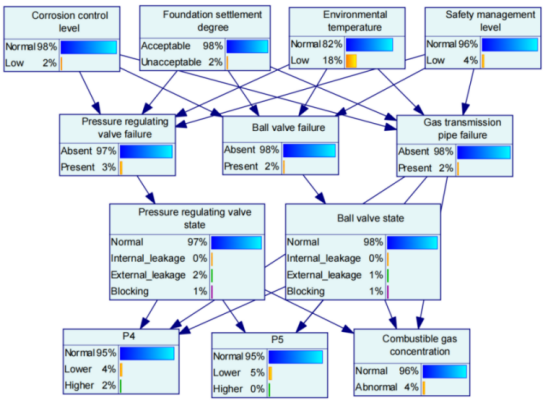


Figure 4.10 BN of object 4

Under normal circumstances, the outlet pressure fluctuates up and down within a specific range. Still, if there is a continuous drop or rise in pressure, the GTS system may have some abnormal situation. Utilizing anomalous monitoring data and combustible gas concentration in each object area can be initially determined which object is most likely to have potential risk.

Usually, the emergency shut-down valves at the inlet and outlet stations are fully open. And when the outlet pressure is abnormal, it should be checked for faults in the first place. Therefore, we do not consider objects 1, 9 and 10 when diagnosing faults in the system. The OOBN model for a GTS system used for diagnosis is shown in Figure 4.11.

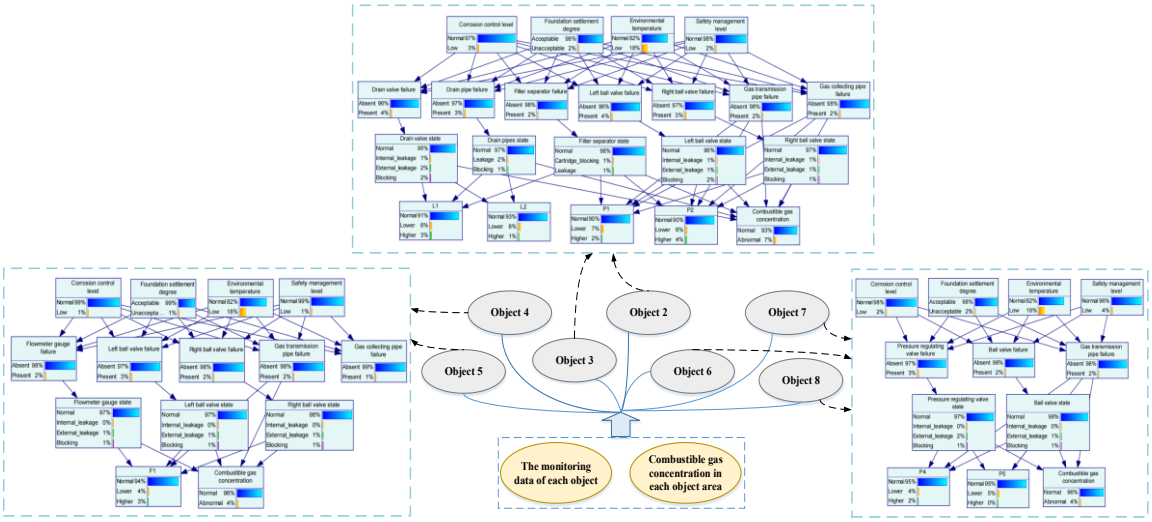


Figure 4.11 OOBN model

4.4 Discussions

4.4.1 Model performance analysis

Historical failure records of the GTS are used to evaluate the performance of the proposed model. To better illustrate the ability of the PAA-CUSUM algorithm to detect potential system risks with its sensitive detection, we selected data from 0 to 24 hours before the failure record to conduct the analysis. The related Bayesian inference of GTS

can be divided into three steps.

First, the PAA-CUSUM algorithm detects an abnormality in the outlet pressure over 24 hours. Then, by calculating the change degree of indicator readings in each object area, the object with the highest potential risk can be determined with the most significant change degree.

Then, determine whether the abnormal indicator reading shows a lower or higher reading than usual. For example, if it offers a downward trend, the state of the pressure node can be set to "Lower," and the value of Δp can be calculated. If a component failure node in the obtained object has a Δp that is significantly higher than the others, the component has the highest potential risk. If it is not that significant, according to Δp, a ranking list of the suspected objects can also be provided. Finally, the component state node is checked, where the state with the most significant Δp is the possible failure state of the component.

The data has already been tested in phase I, and no out-of-control occurs, so we can directly use the PAA-CUSUM chart to capture a very small shift.

Case 1:

Due to impurities such as water and hydrogen sulfide in the gas, corrosion leakage occurs in the pipe located between BV and PRV in object 4. The PAA-CUSUM algorithm detected the outlet #1 pressure as a continuous drop, as shown in Figure 4.12, indicating somewhere abnormal in the system. Therefore, the relevant monitoring data of each object was retrieved for PAA-CUSUM analysis, and the change degrees were calculated using Eq. (4.16)-(4.17). Figure 4.13 shows the indicator readings based on the PAA-CUSUM algorithm in objects 4 and 7. The other indicator readings are 0 or fluctuating around 0 after CUSUM analysis, which illustrates no risks in the object area. The calculated change degree of objects 4 and 7 are 2.21 and 1.27, respectively. Therefore, object 4 has the highest potential risk.

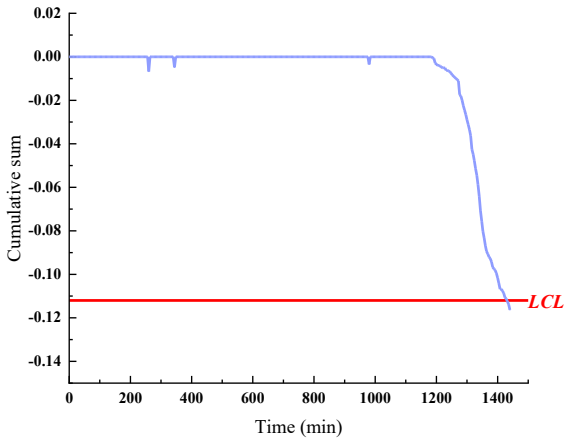
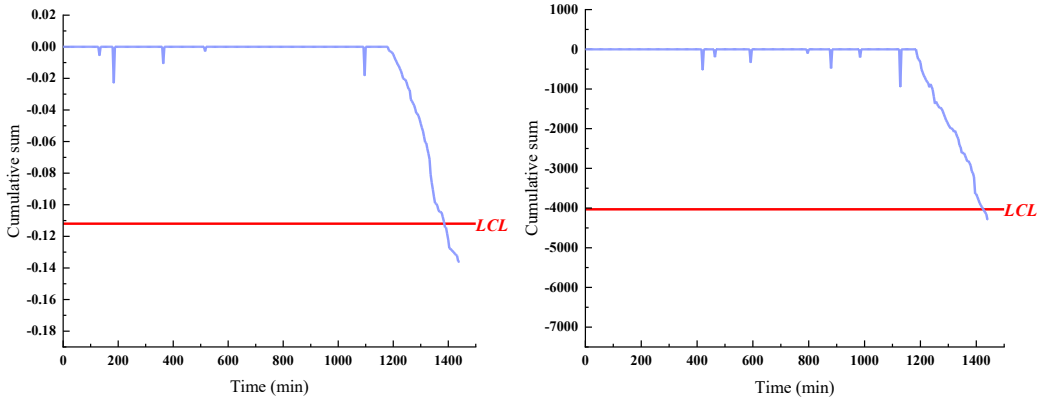


Figure 4.12 CUSUM chart of outlet #1 pressure



(a) Object 4 (b) Object 7

Figure 4.13 CUSUM chart of abnormal objects

Combustible gas concentration measurement is an additional way to determine the risk of leakage in the system. Based on Figure 4.10, the two pressure nodes are set to "Lower," and the combustible gas concentration node is set to "Abnormal." The difference Δp between the calculated prior probabilities and the posterior probabilities is shown in Figure 4.14.

According to Figure 4.14, the most likely failure in object 4 is the gas transmission pipeline leakage because of its highest Δp value. The subsequent most likely failure is the ball valve with external leakage. The most likely risk factor is corrosion. According to the failure records, it is due to corrosion that causes gas pipeline leakage, thus indicating the correctness of the proposed method.

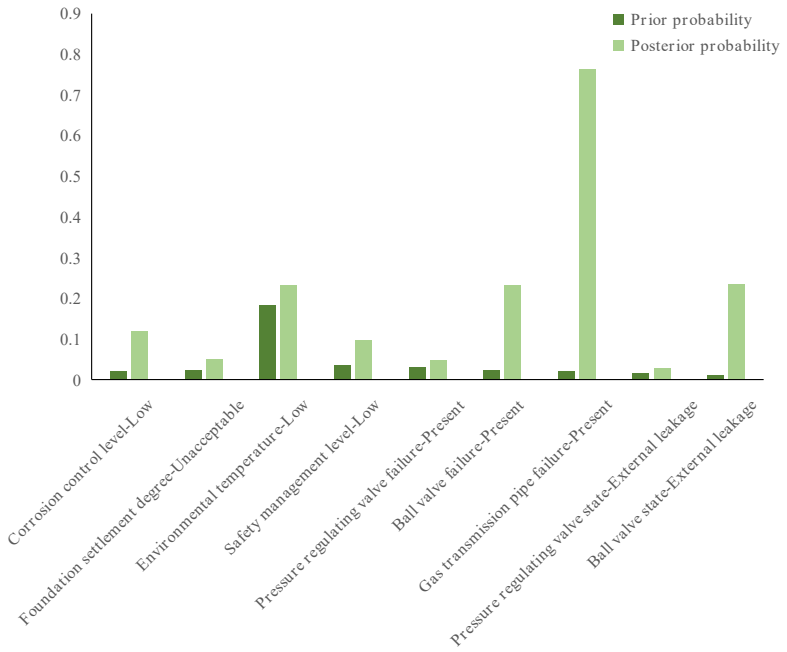


Figure 4.14 Prior and posterior probabilities

Case 2:

Due to the cold winter weather, the flowmeter gauge located in object 6 was frozen and blocked. The PAA-CUSUM algorithm detected the outlet #2 pressure as a continuous increase, as shown in Figure 4.15, indicating abnormal in the system. Therefore, the relevant monitoring data of each object was retrieved for PAA-CUSUM analysis. Figure 4.16 shows the indicator readings based on the PAA-CUSUM in object 6. The other indicator readings are 0 or fluctuating around 0 after CUSUM analysis, which illustrates no anomalies in the object area.

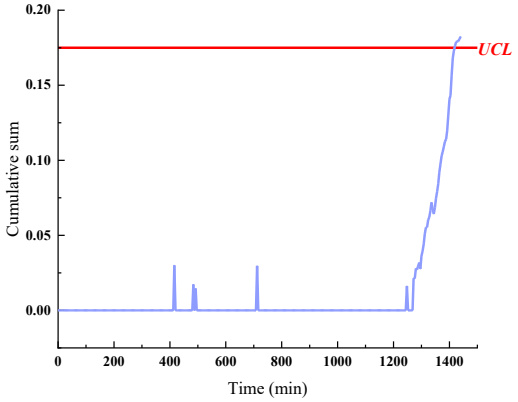


Figure 4.15 CUSUM chart of outlet #2 pressure

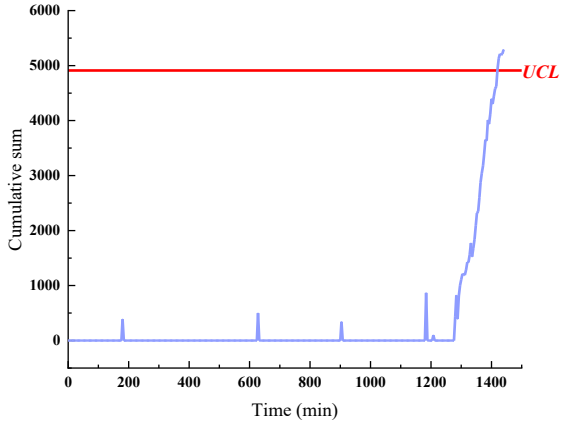


Figure 4.16 CUSUM chart of object 6

Based on Figure 4.9, the leaf node about pressure in object 6 is set to “Lower.” The difference Δp between the calculated prior probabilities and the posterior probabilities is shown in Figure 4.17.

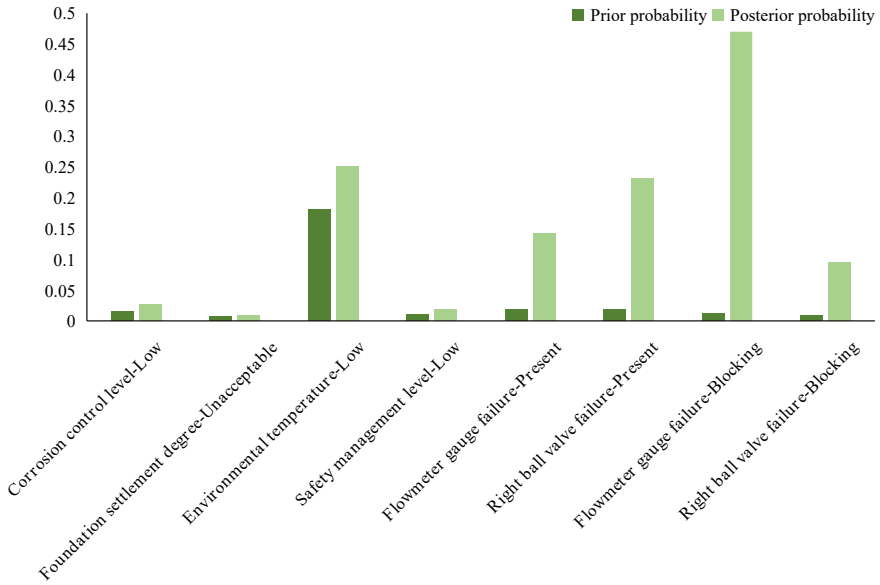


Figure 4.17 Prior and posterior probabilities

According to Figure 4.17, the most likely failure in object 6 is the flowmeter gauge blocking because of its highest Δp value. The subsequent most likely failure is the ball valve with external leakage. The most likely cause of failure is low temperature. According to the failure records, it is due to low temperature that causes flowmeter gauge blocking, thus indicating the correctness of the proposed method.

Similarly, we identify other precursors, as shown in Table 4.6.

Table 4.6 Other cases

Case number	Object with maximum potential risk	Component with maximum potential risk	Actual failure record
1	Object 2	LBV(internal leakage, Δp=0.5113) RBV(internal leakage, Δp=0.2626)	Object 2, RBV, internal leakage
2	Object 7	LBV(blocking, Δp=0.4462)	Object 7, LBV, blocking
3	Object 3	FS（blocking, Δp=0.3219）	Object 3, FS, blocking

The case studies above demonstrate the effectiveness of the proposed method in detecting anomalies and identifying accident precursors before formal reporting fails. According to case#1 in Table 4.6, the component that fails may not be the one with the highest potential risk, but it must have a high ranking of potential risk. Among the conclusions drawn from the five cases above, the components with the highest potential risk matched 80% of the historical records, and the top two components with potential risk matched 100% of the historical records. In practice, GTS operators should check the components with a high order of potential risk in objects with accident precursors to prevent multiple parts from failing simultaneously.

4.4.2 Sensitivity analysis

The volume of information communicated between two or more variables is called mutual information. Take two variables as an example; the stronger the correlation between the variables, the larger the mutual information. Eq. (4.18) may be used to calculate the mutual information between Y and X [254].

$$\mathrm{H(X:Y)} = \sum_{y\in Y}\sum_{x\in X} \mathrm{P(x,y)log}\left(\frac{\mathrm{P(x,y)}}{\mathrm{p(x)p(y)}}\right) \tag{4.18}$$

where $P(x,y)$ represents the joint probability distribution of X and Y, $p(x)$ and $p(y)$ represent marginal distribution function of X and Y.

Take the gas transmission pipe in object 4 as an example; the calculated mutual information after normalization is shown in Table 4.7. To verify the correctness of the results obtained by mutual information calculation, the sensitivity analysis function of GeNIe software is utilized, as shown in Figure 4.18. The gray nodes indicate irrelevance, and the darker the red nodes, the higher the sensitivity, which is consistent with the

results of mutual information calculation. Therefore, for the gas transmission pipe in object 4, the risk factors are as follows: corrosion>safety management>foundation settlement>environmental temperature.

Table 4.7 Mutual information between gas transmission pipe and first layer nodes in object 4

Node	Mutual information(%)
Gas transmission pipe	100
Corrosion control level	4.09
Safety management level	1.21
Foundation settlement degree	0.31
Environmental temperature	0.19

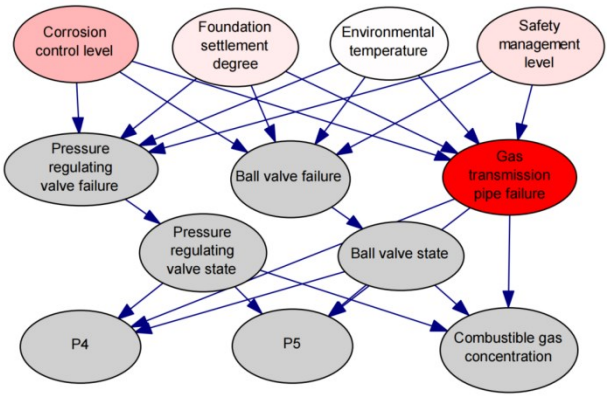


Figure 4.18 Sensitivity analysis of object 4

Similarly, the critical risk factors of other components in the GTS system are identified and summarized, as shown in Figure 4.19. It can be concluded that corrosion is the risk factor that has the most significant impact on this GTS system.

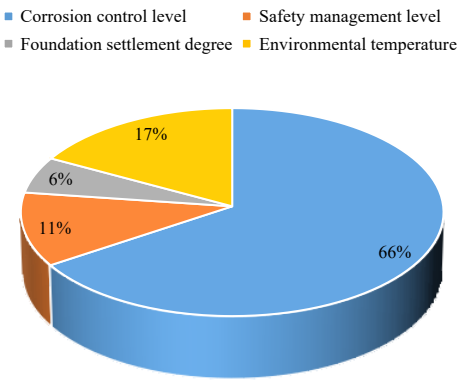


Figure 4.19 Critical risk factors of components in the GTS system

4.4.3 Model comparison

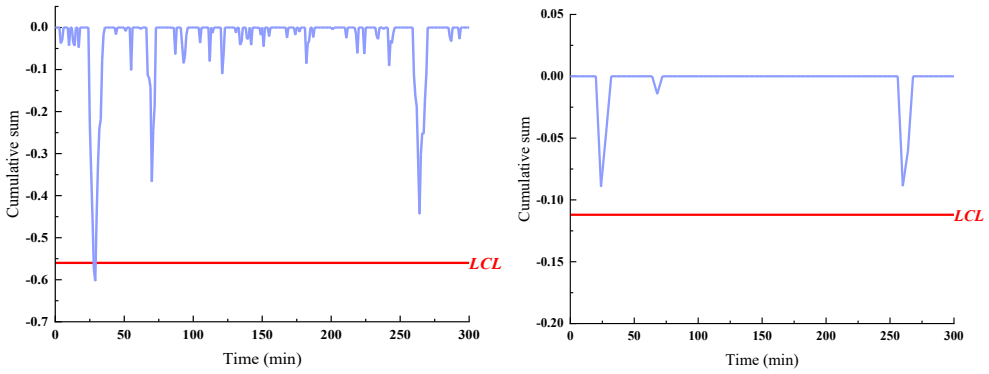
Table 4.8 shows the basic parameters of the different modeling methods. Compared to BN, OOBN with LNG reduces the number of nodes by over 50% and the number of CPTs by over 60%. Compared to OOBN, OOBN with LNG also reduces the number of CPTs by over 20%. Therefore, OOBN with LNG is beneficial in terms of node number and CPT input, which can simplify the model and minimize the GTS operators' workload, allowing users to operate more efficiently.

Table 4.8 Comparison of different methods

Model parameters	BN	OOBN	OOBN with LNG
Node number	106	47	47
CPT number	976	452	348
Highest Δp value in case 1	72.35%	72.35%	74.27%
Highest Δp value in case 2	44.28%	44.28%	45.70%

In case 1 and case 2, the risk rankings calculated using BN and OOBN are the same as those calculated using OOBN with LNG, indicating the accuracy of OOBN with LNG is guaranteed. The Δp value measures the difference between the posterior and prior probability of BN. A larger Δp value indicates that the model is more sensitive to the inputs and better able to detect potential risk factors. As a result, OOBN with LNG modeling is more accurate.

To demonstrate the applicability of the proposed PAA-CUSUM method, we randomly intercepted 300 minutes of normal outlet pressure data for analysis. As shown in Figure 4.20, the red line indicates the alert line. After applying the CUSUM algorithm to the phase II data, it causes false alarms. But we discover that by reducing the data dimensionality, the PAA-CUSUM algorithm may effectively minimize the false alarm rate and create a decent denoising impact.



(a) CUSUM (b) PAA-CUSUM

Figure 4.20 Model comparison

4.4.4 Model validation

The three axioms proposed by Jones et al.[255] are used to perform model validation. If the model is robust, the rise or reduction in each parent node's prior probability should result in a proportionate increase or decrease in the child node's posterior probabilities. The evidence set's impact should always be stronger than the sub-evidence's impact. According to the sensitivity analysis, the results satisfy the axioms, partially demonstrating the proposed model is rational and valid.

4.5 Conclusions

This research aims to present a novel method for risk analysis in GTS systems. This chapter establishes a risk analysis model framework based on OOBN according to the structure and process flow of the GTS system, which can effectively identify the system's accident precursors. For the components, the analysis result can provide the ranking of potential risk from high to low, as well as their possible failure types and failure causes, which can guide the GTS operator to perform efficient risk rehabilitation.

In this chapter, the structure of BN network segments of each object is established through the mapping of FMEA. The CPT parameters of each BN segment are determined through the combination of LNG and EM algorithms, which increases the expression of uncertainty and solves the problems of small data samples and imperfect

historical records for the modeling process of engineering practice. This study uses the PAA-CUSUM algorithm to analyze the monitoring data and then inputs the results into the BN segments of each object for risk analysis. The PAA algorithm performs well in terms of data dimensionality reduction and denoising, while the CUSUM algorithm performs well in terms of identifying small trend changes in the data.

The proposed method is a promising approach to risk analysis of the GTS system. However, the followings are some limitations. (1) The risk factor classification can be more thorough. With comprehensive data, more in-depth analysis may be conducted in the future. (2) Only the daily regular operation of the GTS system is considered, but not the risks under other operating conditions such as venting. (3) Future studies on more complicated systems will require a more precise component separation, such as the process piping segregated according to pipe diameter and other factors. As a result, future studies should concentrate on these issues to ensure the stable and smooth operation of the GTS system.

Chapter 5: Reliability analysis of corroded pipes using MFL signals and Residual Neural Networks

5.1 Introduction

Pipelines constructed of carbon steel are used to transport oil and natural gas across long distances to serve the world's energy needs. As most of the pipes are buried underground, it is possible for the soil and other environmental factors that influence pipelines to induce various types of corrosion [256, 257]. These types of corrosion threaten the pipeline's integrity, considerably increase the hazards associated with pipeline operation, and may even lead to accidents [258].

According to the PHMSA [259], the majority of ILI detectable leaks are caused by corrosion in high-consequence areas (HCA), as shown in Figure 5.1. In addition, according to [260], over 60% of reported pipeline failures in southern Mexico may be attributed to external corrosion. These statistics illustrate that corrosion is one of the most significant threats to pipeline integrity. Therefore, pipeline operators should prioritize dealing with pipeline corrosion and analyzing the pipeline's reliability to ensure safe operation.

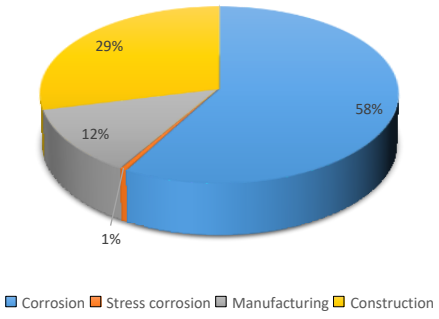


Figure 5.1 Gas transmission leak cause of HCA in the 2005 - 2022 period

Pipeline ILI carried out regularly can effectively reduce the likelihood of incidents and allow the operators to comprehensively understand the pipeline's corrosion

conditions and reliability status, thereby taking prompt action to maintain and repair the pipeline's defects. For example, Xie and Tian. [261] and Abubakirov et al. [262] utilized the ILI results to find the optimal inspection interval. Furthermore, based on the ILI data, Amaya-Gómez et al. [263] proposed a dynamic segmentation and clustering to analyze the integrity of corroded pipeline; Wang et al. [264] established a stochastic defect growth model for corroded pipelines.

Magnetic flux leakage (MFL) is the most common technique in pipeline ILI [265]. Therefore, more and more scholars have studied the MFL signals to obtain defect information. Examples include using pattern-adapted wavelets and artificial neural networks (ANN) to estimate the length and depth of metal-loss defects [86]; using a multiscale Single Shot MultiBox Detector (SSD) network to automatically identify the location of girth weld, spiral weld, and defect [266]; using improved particle swarm optimization and a RBFNN to recognize the width and depth of defects [267].

Based on the obtained defect information, a subsequent reliability analysis can be conducted. In-service pipelines experience degradation as a result of various risk factors. The degradation process is contingent upon time, leading to the temporal variability of the structural reliability of pipelines. Therefore, it is necessary to analyze pipelines' reliability and future reliability trends to better conduct integrity management, optimize costs, and derive the optimal re-assessment time interval [145, 268]. The related studies are shown in Table 5.1.

The following conclusions may be drawn from the research in Table 5.1: (1) In light of the development of "Oil and Gas 4.0," machine learning and deep learning have significant promise in the applications of the pipeline industry [269], particularly in terms of data analysis and interpretation. (2) Some research only considers the defect depth and length without considering the defect shape or defect's effective area when establishing the algorithm models, leading to less precise reliability analysis results. (3) Most research focuses on processing MFL inspection data to obtain the defect's geometry and then using the geometry information to conduct the subsequent analysis, including reliability prediction, failure pressure prediction, residual strength prediction,

et al. The calculation and analysis procedures are complicated and time-consuming. In this way, the pipeline operators may take much longer to obtain the final reliability result rather than evaluating the pipeline status efficiently and taking maintenance measures promptly.

Table 5.1 Some relevant studies

Literature	Proposed method	Problems solved
[136]	Back propagation neural network method	Predict reliability of corroded pipeline under combined loadings
[270]	Multilayer perceptron and modified feedforward neural network	Predict residual strength of corroded pipelines
[271]	First-order reliability method and limit state function	Analyze the reliability of submarine pipeline with corrosion defects
[272]	FE-based model	Analyze the defect growth and predict failure pressure
[273]	Extreme value analysis	Predict the depth of extreme pits
[137]	Back propagation neural network	Predict triple failure pressure of corroded pipes
[274]	Subset simulation	Analyze the reliability of corroded natural gas pipeline

This chapter proposes an approach for directly obtaining pipelines' reliability and provides pipeline operators with risk management suggestions using MFL inspection data. Compared to previous research, the contributions are summarized as follows. (1) The FE simulation of MFL signals and reliability analysis are combined through deep learning methods, allowing the MFL signals to be directly mapped to the pipeline's reliability. The principles of reliability calculation based on the effective area model rather than just depth are also effectively integrated into this mapping process, which can obtain more accurate results. (2) A novel ResNet-based reliability prediction method is proposed, making the analysis process more accurate and efficient. The case in this chapter shows that, compared to traditional methods, the proposed model's accuracy is more than 20% higher, and the computational efficiency has been increased by 200 times. (3) The re-assessment interval optimization method is also integrated into the deep learning algorithm, which can automatically generate the optimal time interval

and can effectively help pipeline operators make decisions based on the MFL inspection signals.

The rest of this chapter is structured as follows. Section 2 introduces some basic concepts and the proposed method; Section 3 discusses the model performance; Section 4 provides an industrial application to demonstrate the proposed method; Section 5 presents the conclusions.

5.2 Methodology

5.2.1 ResNet

A convolutional neural network (CNN) is a feedforward neural network with a deep structure and convolutional computation [275]. CNN has been demonstrated to deal effectively with audio, time series, and signal data [276, 277]. However, the accuracy of the CNN network reaches a saturation point or possibly drops when the number of CNN layers comes to a certain level, making it challenging to train the model. He et al. [278] introduced a residual learning framework (residual network, ResNet) to solve the degradation problem that arose in deep networks and demonstrated that the framework could be easily trained even with deep architectures.

A typical CNN block and a ResNet block are depicted in Figure 5.2(a) and (b), respectively. As can be found, the ResNet block differs from the standard CNN block in having one additional shortcut connection [279]. Shortcut connections guarantee that the gradient may still be propagated backward even after many neural network layers. This way, the neural network model's performance and training effectiveness are significantly enhanced [280].

The residual network is composed of a series of residual blocks. A residual block can be expressed as:

$$H = X + F(X) \tag{5.1}$$

where H denotes the output, X represents the identity mapping and $F(X)$ represents the residual mapping. The specific structure is shown in the Figure 5.2.

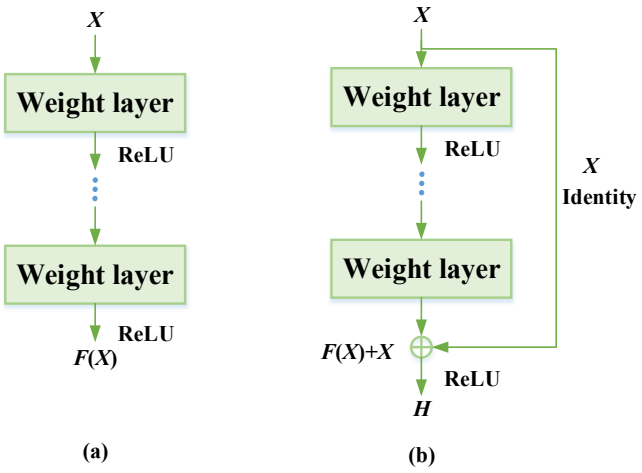


Figure 5.2 Structure comparison of standard CNN block and ResNet block

5.2.2 The proposed method

The specific steps of the proposed method are shown in Figure 5.3. First, the three-axial MFL inspection data is obtained from the FE simulation. Then, the reliability function is established using the limit state function (LSF) and MCS. According to the defect depth and effective area model, the LSF is established. After reliability calculations, the data set is labeled. In the algorithm model's architecture constructing stage, the general residual blocks are first constructed, and then specific ResNet architectures are constructed. In this chapter, the structures of 4th-Runge–Kutta Net (RK4-Net) and Euler-Net are integrated into the model establishment. After the training and testing, the model can be applied to the industry analysis and provide managerial suggestions for the decision-makers.

This study proposes an approach for quickly determining pipelines' reliability, optimal re-assessment interval, and cost rate (CR) based on MFL inspection data and ResNet. The fundamental concept of ResNet is to establish an "identity shortcut connection," which may boost both the performance of the model and the efficiency of its training. Compared with the methods used in the existing studies in Table 5.1, ResNet's image processing abilities are superior, allowing it to more effectively extract the characteristics of MFL signals, ultimately leading to more satisfying results. In establishing the LSF and labeling the dataset, the effective area of the defect is also

considered instead of just the defect depth, which leads to more accurate results. In developing the structure of the ResNet block, this chapter refers to the idea of the ordinary differential equation (ODE) solver. It introduces a structure based on the RK4 method, improving the model's accuracy. In addition, the FE model is combined with ResNet, and the large volume of data obtained from FE simulation in this chapter provides good support for ResNet training. Furthermore, the re-assessment interval optimization algorithm is also combined with ResNet, allowing pipeline operators to obtain the optimal time interval for re-assessment and make optimal cost decisions.

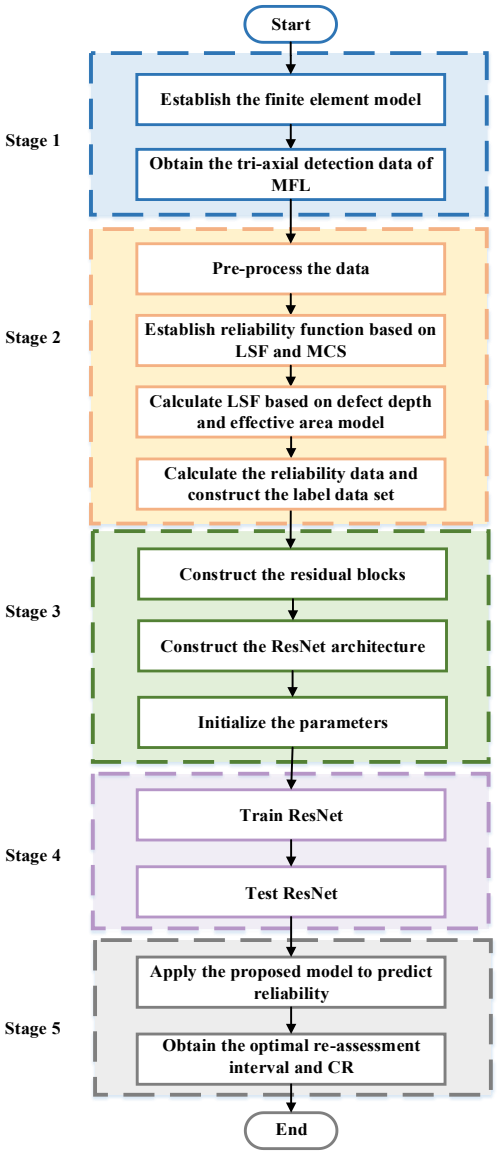


Figure 5.3 The specific steps of the proposed method

5.2.2.1 Finite element model

As shown in Figure 5.4, when performing an MFL inspection, a permanent magnet is used to magnetize the ferromagnetic pipe wall almost to the point of saturation. The existence of a defect induces an increase in the flux density in the region immediately around the defect. This, in turn, causes an increase in reluctance and results in the leaking of flux lines into the region around it. Due to this, a magnetic field will begin to "leak," which may be detected using Hall sensors [281, 282]. Figure 5(a) shows a defect sample in a rectangular cuboid shape. Figure 5.5(b) shows an MFL tool for the pipe with an outer diameter of 323 *mm* and a wall thickness of 8 *mm* in the engineering application.

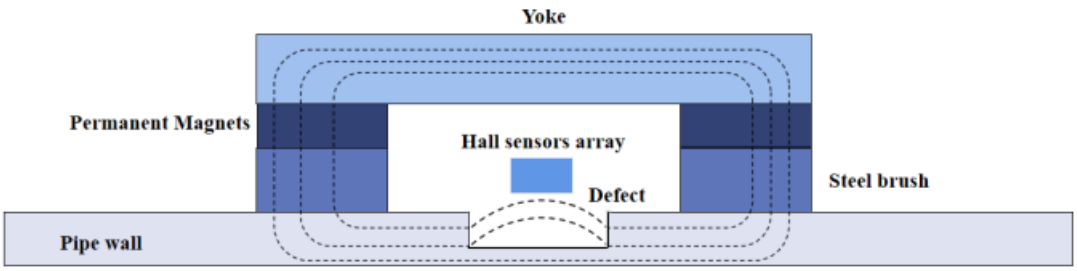


Figure 5.4 The diagram of MFL inspection principle

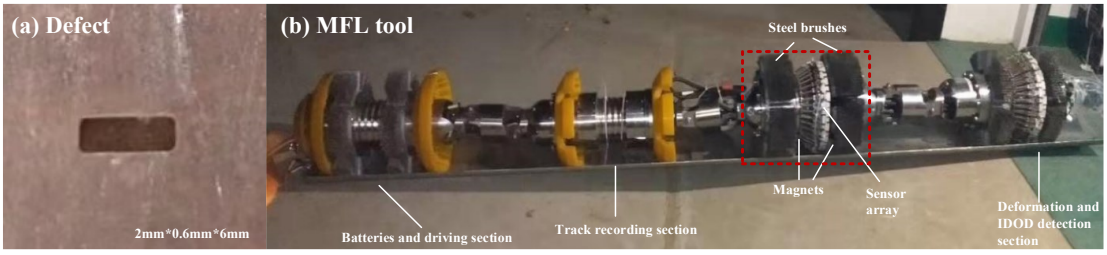


Figure 5.5 The industrial situation

In this chapter, a three-dimensional solid equivalent simplified model based on the MFL inspection principle and the industrial situation is established by the FE method, as shown in Figure 5.6, which can study the leakage magnetic field of pipeline defects and obtain a sample database for analyzing the reliability of pipe sections. Furthermore, this model may be used to determine the three-axial MFL signal at the defect as well as the spatial distribution of the leakage magnetic field, both of which can be used as the

basis for reliability evaluation. In practical engineering, MFL detection requires significant time and financial support to obtain rich data, which is extremely challenging to perform. The FE simulation not only substantially reduces the time cycle, but it also achieves the desired statistical results by approaching as close to the actual condition as possible.

In Figure 5.6, the coercivity of the magnet is set as 6.6×10^5 A/m, and the conductivity of steel is set as 2×10^6 S/m. The pipe material is steel Q235, the outer diameter is 323*mm*, and the wall thickness is 8*mm*. COMSOL Multiphysics is employed to carry out the analysis. The solid model is divided into triangular elements to generate the mesh model. Since the model's overall size is much larger than the size of the defects, the mesh in the neighboring area of the defects is encrypted to obtain more reasonable results while ensuring computational efficiency. The B-H curve of the components is shown in Figure 5.7. The lift-off distance is set as 2*mm*. The distance of the defect along the axial direction of the pipe is defined as the defect length; the distance of the defect along the circumferential direction of the pipe is defined as the defect width; the distance of the defect along the radial direction of the pipe is defined as the defect depth. The sample MFL signal is shown in Figure 5.8.

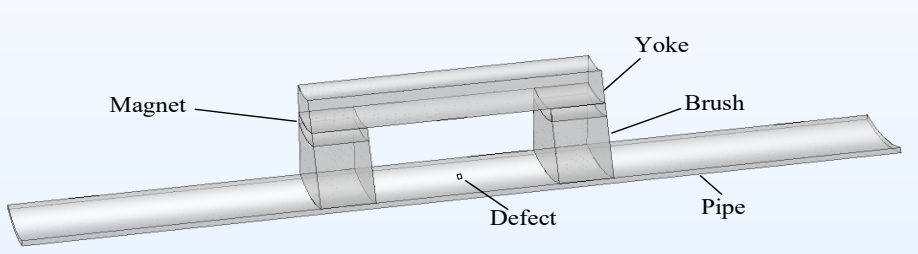
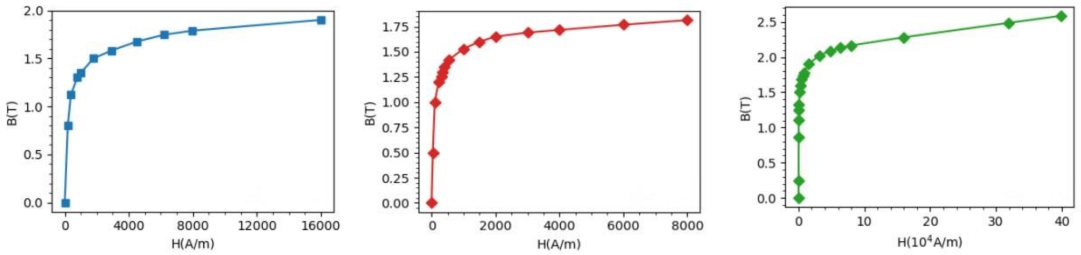
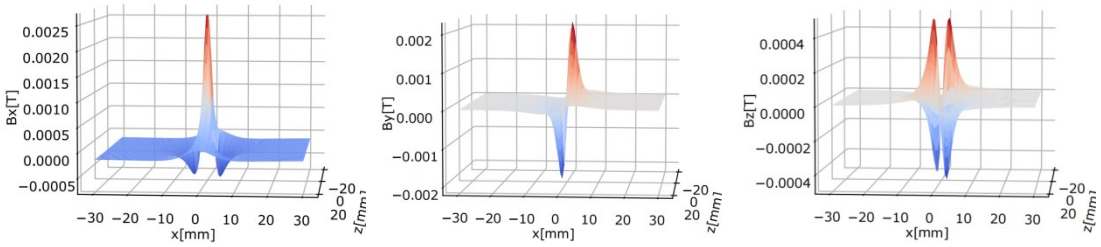


Figure 5.6 The FE model



(a) B-H curve of the pipe (b) B-H curve of the yoke (c) B-H curve of the brush

Figure 5.7 B-H curve of the components



(d) Bx (e) By (f) Bz

Figure 5.8 MFL signal sample

We modeled the MFL signals when detecting the cuboid and ellipsoidal defects. Through the FE simulation, we randomly generated 3750 sets of data, including 1750 sets of ellipsoidal defect data (diameter in the range of 2*mm* - 20*mm*, depth in the range of 0.5*mm* - 6*mm*) and 2000 sets of cuboid defect data (length and width in the range of 2*mm* - 20*mm*, depth in the range of 0.5*mm* - 6*mm*).

5.2.2.2 Reliability function establishment

The radial depth and axial length may represent pipe corrosion. According to Eq. (5.2) - (5.3), corrosion depth and length are linear factors of time for the corrosion process.

$$d(T) = d_0 + v_r T \tag{5.2}$$

$$L(T) = L_0 + v_a T \tag{5.3}$$

where d_0 and L_0 are the initial corrosion depth and length of the defect, v_r is the radial corrosion rate, v_a is the axial corrosion rate, and T is the time elapsed since the last internal inspection.

For an operational pipeline subject to a risk brought on by corrosion, perforation leaking and bursting are two potential failure mechanisms [27]. This chapter mainly considers the single corrosion defect. The maximum allowable corrosion depth of the pipe cannot exceed 80% of the pipe wall thickness [283]. Hence, the difference between the maximum allowable corrosion depth and the actual corrosion depth may be used to define the LSF of pipe corrosion perforation, as shown in Eq. (5.4).

$$g_1(T) = 0.8t - d(T) \tag{5.4}$$

where t is the wall thickness of the pipe, d indicates the depth of the corrosion defect.

According to the pipe burst failure mechanism, for developing corrosion defects,

the LSF of burst failure is defined as the difference between the pipe failure pressure p_f and the actual operating pressure p_o, as shown in Eq. (5.5).

$$g_2(T) = p_f(T) - p_0(T) \tag{5.5}$$

The reliability calculation is shown as follows:

$$R(T) = 1 - P_f(T) = 1 - P[g(T) < 0] \tag{5.6}$$

where P_f (T) is the total failure probability of the two failure types, which can be calculated by the MCS method. It has proven to be an effective tool for the reliability evaluation of pipe sections with corrosion defects. Its specific process can be referred to [284]. Typically, to attain the requisite precision, the number of random variable parameters sampled is 10^6 times. And the detailed probabilistic characteristics of the random variables involved in the reliability analysis can be found in [261], the variables are assumed to follow normal distributions. The specific data used for case study is provided by pipeline operators. The operating pressure is 6.7 MPa with a standard deviation of 0.67 MPa, the radial corrosion growth rate is 0.3 *mm*/year with a standard deviation of 0.03 *mm*, and the axial corrosion growth rate is 8 *mm*/year with a standard deviation of 0.5 *mm*.

5.2.2.3 Effective area model

This chapter uses PRCI RSTRENG (effective area model) [285, 286] to determine the failure pressure p_f in the LSF, as shown in Eq. (5.7) - (5.9).

$$p_f = \frac{\sigma_{flow} 2t}{D}\left[\frac{1-(A(T)/A_0)}{1-(A(T)/A_0)M^{-1}}\right] \tag{5.7}$$

$$M = \sqrt{1 + 0.6275\left(\frac{L(T)}{\sqrt{Dt}}\right)^2 - 0.003375\left(\frac{L(T)}{\sqrt{Dt}}\right)^4}, \text{ if } L(T) \le \sqrt{50Dt} \tag{5.8}$$

$$M = 3.3 + 0.032\left(\frac{L(T)}{\sqrt{Dt}}\right)^2, \text{ if } L(T) > \sqrt{50Dt} \tag{5.9}$$

where σ_{flow} is flow stress, D is pipe outer diameter, t is the wall thickness, M is the bulging factor, L is the defect length. A_0 denotes the original cross-section area, as shown in Eq. (5.10), $A(T)$ represents the metal loss area of a complex defect profile in the axial plane, which is the "effective area", the specific calculation formula is shown in Eq. (5.11). Figure 5.9 also demonstrates the effective area of a defect.

$$A_0(T) = tL(T) \tag{5.10}$$

$$A(T) = \alpha L(T)d(T) \tag{5.11}$$

where $\alpha = 1$ for rectangular equivalent shape, $\alpha = 2/3$ for parabolic equivalent shape, $\alpha = 0.85$ for mixed equivalent shape [287].

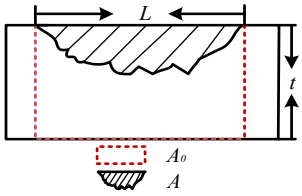


Figure 5.9 Diagram of the effective area *A*

5.2.2.4 ResNet establishment

ResNet is relatively easy to adapt, and accuracy may be raised while increasing the network's depth [288]. The structure of the proposed method includes input and output layers, a convolution block, a residual block, and fully connected layers, as illustrated in Figure 5.10. The input for this study is 100*100-pixel images of the three-axial signal that the MFL tool detected. The output is a one-dimensional matrix representing reliability values for each year over the next 30 years, the optimal re-assessment interval, and the cost rate (CR). According to Figure 5.11, the CNN block developed for this study includes two convolution operations and a max pooling operation. Additionally, each residual block has both mainline and shortcut connections, as shown in Figure 5.12. This chapter uses both Euler-Net and RK4-Net to establish the residual blocks for further comparison.

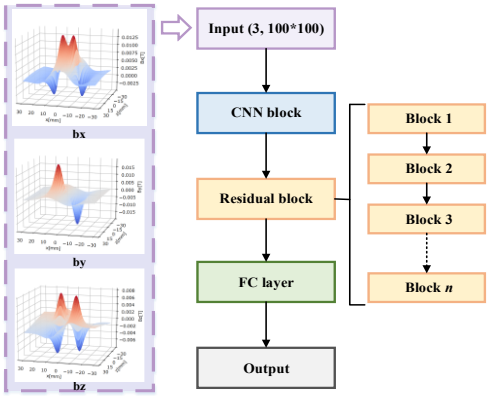


Figure 5.10 The general ResNet structure

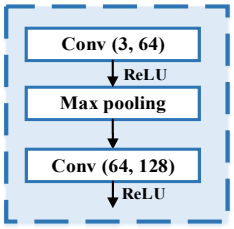


Figure 5.11 The structure of the CNN block

When establishing the ResNet block structure, in addition to building the traditional Euler-Net structure as shown in Figure 5.12(a), this study also draws on the idea of ODE solvers [289, 290] and designs the network structure in the form of a fourth-order Runge-Kutta (RK4) method [291], as shown in Figure 5.12(b). The mathematical expressions are:

$$k_1 = F(X^{in}) \tag{5.12}$$

$$k_2 = F(\frac{1}{2}k_1 + X^{in}) \tag{5.13}$$

$$k_3 = F(\frac{1}{2}k_2 + X^{in}) \tag{5.14}$$

$$k_4 = F(k_3 + X^{in}) \tag{5.15}$$

$$X^{out} = X^{in} + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{5.16}$$

where k represents the computation of each module in the residual block, X^{in} and X^{out} represents the block's input and output, respectively.

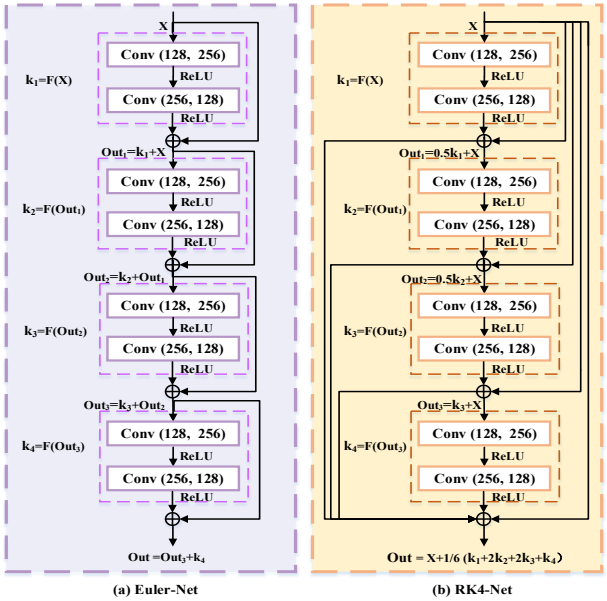


Figure 5.12 The structure of the ResNet block

The typical ResNet expressions and structure reveal that the ResNet constructs complex transformations by assembling a series of transformations to a hidden state. Furthermore, many studies have demonstrated that the typical ResNet's iterative updates may be seen as an Euler discretization of a continuous transformation. In the limit, we may parameterize the continuous dynamics of hidden units using an ODE described by a neural network as we add more layers and make smaller steps. Therefore, the standard ResNet's update formulation can be regarded as the numerical Euler solution of the ODE. For a comprehensive description, see [292]. And as for the numerical analysis, both Euler and RK4 methods are commonly used to solve ODEs. When the step size h is very small, the local truncation error of the RK4 method is on the order of h^5, whereas the total cumulative error is on the order of h^4. However, with the Euler method, each step's mistake is proportional to h^2, and the overall error is proportional to h. As a result, the RK method's accuracy is higher than that of the Euler method. In this study, the conventional Euler-Net is improved to RK4-Net, upgrading the Euler forwards design scheme to a higher-order design scheme, which can increase the model's robustness and accuracy.

5.2.2.5 Training and testing of ResNet

The simulated MFL detection signal data were randomly divided into two sets, the training and test sets. 80% of the samples for each defect shape are randomly selected for training, and the remaining 20% are used to verify the validity of the training model. The training set is used to establish a fitted relationship between the reliability in 30 years and the MFL signal, and the test set is used to evaluate the accuracy of the established fitted relationship. The batch size is set as 32. The epoch is set as 200. The learning rate is set as 10^{-4}. We can obtain a data-driven pipeline reliability prediction model when the established ResNet algorithm is fully trained.

5.2.2.6 Re-assessment interval optimization

Reliability analysis aims to support pipeline operators' risk management and decision-

making processes. Therefore, on the basis of the reliability and its trend obtained from the ResNet model, the pipeline operation and maintenance costs can be optimized and recommendations for re-assessment intervals can be provided.

Each pipe section can be considered a "series system" [261]. Then the failure probability of a pipe section with multiple independent corrosion defects can be expressed as:

$$PF_{pipe} = 1 - \prod_{i=1}^{n} R_{defect,i} \tag{5.17}$$

where PF_{pipe} is the failure probability of the evaluated pipe section, $R_{defect,i}$ represents the reliability of the i-th defect, n represents the number of defects.

To help pipeline operators better understand the consequences intuitively, all of the consequential losses that result from an incident are given a monetary value-based description, as shown in Eq. (5.18) - Eq. (5.19) [293, 294].

$$C_{con,t} = (C_{eco} + C_{env} + C_{hum})/(1 + r)^t \tag{5.18}$$

$$C_{eco} = C_{lp} + C_{dp} + C_{re} \tag{5.19}$$

where $C_{con,t}$ represents the total cost of the failure consequence changing with time t, C_{eco} is the economical cost, C_{env} is the environmental cost, C_{hum} is the cost converted from the damage to human. C_{lp} is the cost of lost production, C_{dp} is the cost of deferred production, C_{re} is the cost of repair, r is the discount rate.

The net present value of total cost is expressed as shown in Eq. (5.20) [295].

$$C_{total,t} = (C_{insp,t} + C_{repl,t} \cdot PF_{pipe} + C_{con,t} \cdot PF_{pipe} + C_{main,t})/(1 + r)^t \tag{5.20}$$

where $C_{total,t}$ represents the total cost changing with time t, $C_{insp,t}$ is the cost from ILI, $C_{repl,t}$ is the replacement cost, $C_{main,t}$ is the maintenance cost.

The cost rate regarding the failure probability is expressed as:

$$CR(PF_{th}) = \frac{C_{total,t}(PF_{th})}{T(PF_{th})} \tag{5.21}$$

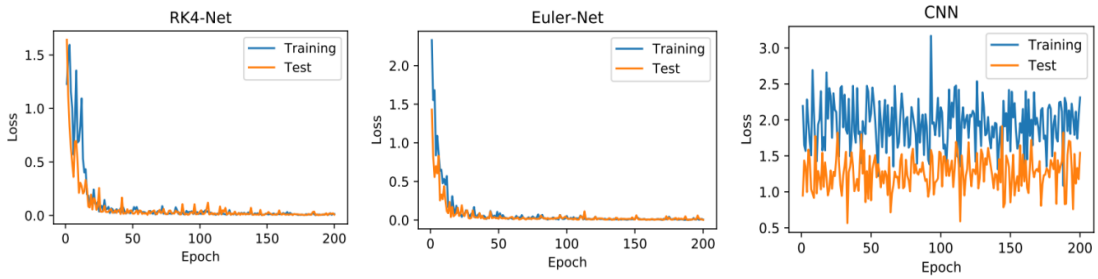
$$T(PF_{th}) = inv(\mathbb{I} - R)(PF_{th}; \gamma) \tag{5.22}$$

where PF_{th} represents the threshold of failure probability, T is the re-assessment time interval, R is the predicted reliability, γ represents the degradation coefficient, $\mathbb{I}$ represents the identity operator.

5.3 Results and discussions

5.3.1 Comparison of different deep learning methods

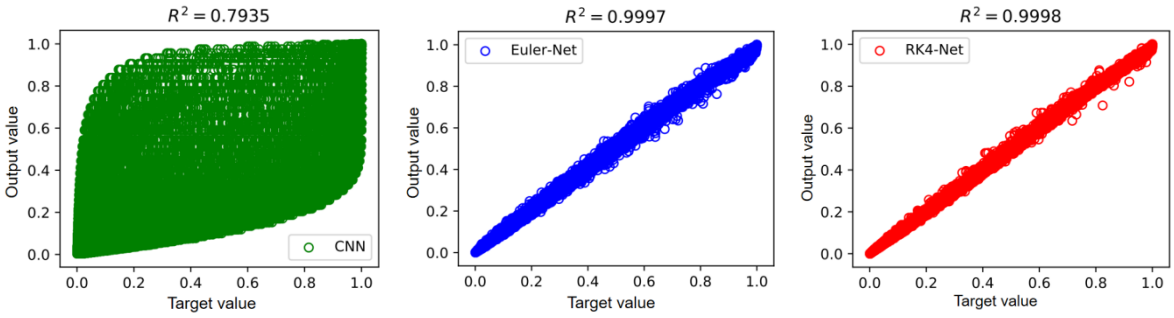
This chapter uses mean square error (MSE) as the loss function to estimate the model's accuracy. As shown in Figure 5.13 (a) and (b), after 200 epochs of training, the desired accuracy can be obtained. In addition, it can also be seen that with the same number of iterations, the CNN fails to converge, indicating that the model performance of the ResNet family is much better than the CNN. It should be noted that the CNN model used for comparison has the same structure as the proposed model except for the shortcut connections.



(a) (b) (c)

Figure 5.13 Iterative process of three structures

The coefficient of determination (R^2) and Root Mean Squared Error (RMSE) are also used to describe the model's performance. The RMSE of CNN, RK4-Net, and Euler-Net are 108.8890, 0.0296, and 0.0571, respectively. Moreover, as shown in Figure 5.14, the coefficient of determination of Euler-Net and RK4-Net are 0.9997 and 0.9998, respectively, which is much better than that of CNN, indicating that the ResNet family performs better. Besides, in the ResNet family, the accuracy of RK4-Net is slightly higher than that of Euler-Net, which may also be explained from the perspective of ODE: the RK4 method has a higher order of error than the Euler method, so the prediction result of RK4-Net is relatively more accurate.



(a) (b) (c)

Figure 5.14 Comparison of the training results of the three models

Nine defect samples are randomly extracted to observe their reliability prediction results using different algorithms, as shown in Figure 5.15. Although the effectiveness of CNN has been demonstrated in the case of small sample sizes and simple models, when faced with more complex situations and more extensive data volumes, the prediction results of CNN are significantly less satisfactory than those of the ResNet family. Furthermore, since deep CNN networks suffer from the degradation problem, the network accuracy saturates and even decreases when the network depth increases. But identity mapping in the ResNet family makes the deep network perform at least as well as the shallow network without degradation problems. Therefore, the accuracy of the ResNet is at least "guaranteed" and more suitable for dealing with massive amounts of MFL inspection data.

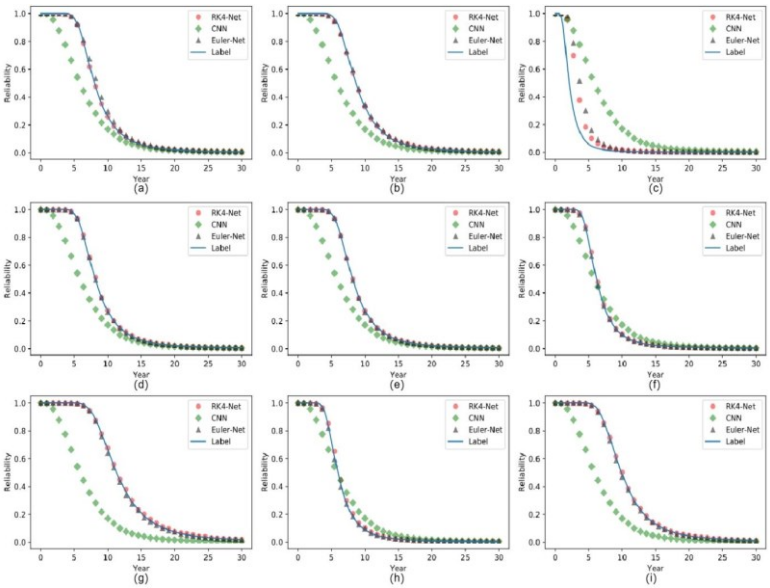


Figure 5.15 Reliability prediction results for nine randomly selected defects

5.3.2 Comparison of traditional and proposed method

When calculating the re-assessment time interval by MCS and RK4-Net from 100 randomly chosen points, we found that the relative error was 0.03%, which is negligible. Therefore, from the analysis above, it can be concluded that the ResNet-based predictions and the results obtained based on MCS are consistent, the results of the proposed method are correct. On this basis, we analyze the time required to obtain the results using different methods in the same operating environment, as shown in Table 5.2. To avoid some extreme cases and for a fair comparison, we did 100 tests for each method and averaged. Each method is used to analyze the reliability of the same defect and the time consumed for each test.

Table 5.2 Comparison of the computing time

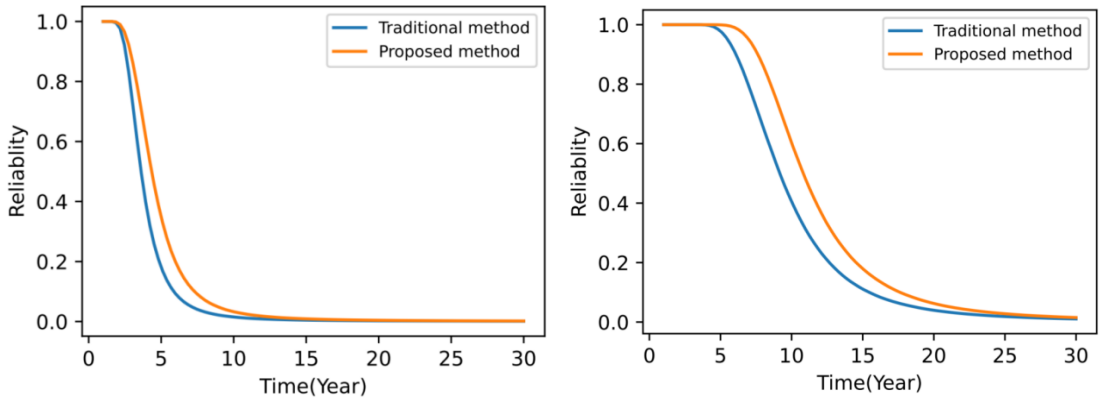
Model	Time(s)
RK4-Net	0.0534 ± 0.0072
Euler-Net	0.0444 ± 0.0041
MCS	10.8357 ± 0.0332

According to Table 5.2, Euler-Net takes less time than RK4-Net due to its more straightforward network structure. However, either Euler-Net or RK4-Net takes far less time than MCS. The ResNet family computes nearly 200 times faster than MCS, demonstrating superior efficiency.

In addition, according to Table 5.2 and Figure 5.14, the accuracy of RK4-Net is slightly higher than Euler-Net, but RK4-Net takes marginally longer to obtain the results. Therefore, in practice, when the situation is more complex and more accurate results are needed, RK4-Net is recommended. On the contrary, if the pipeline operator places a greater emphasis on efficiency, Euler-Net is the more appropriate method to use in cases where a certain degree of accuracy can be guaranteed.

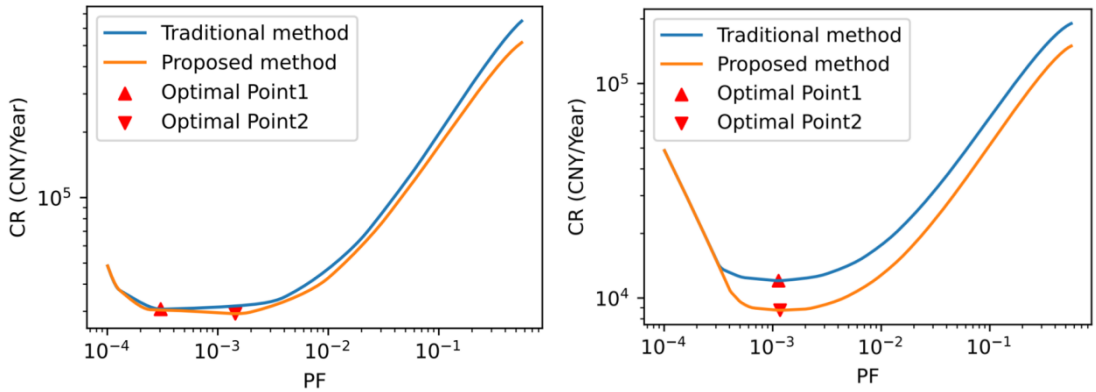
In the traditional method, people usually idealize the defect cross-section as a rectangle [97, 136, 137, 296-298] or only use defect depth and length to obtain reliability [116, 284, 299]. However, this study substitutes the effective area of defects instead of the idealized defect depth and length into the LSF's burst failure calculation, which is more accurate than traditional methods.

We randomly chose two defects for analysis. Figure 5.16 compares the reliability results obtained using the traditional method and the proposed method, showing that the results obtained using the traditional method are more conservative. Figure 5.17 shows the comparison of the CR obtained by the traditional method and the proposed method. For defect 1, the proposed method yields a 6% reduction in CR compared to the traditional method. For defect 2, the proposed method yields a 30% reduction in CR compared to the traditional method. When idealizing defects, a portion of the area that is not a defect is also accounted for, thus leading to more conservative results from traditional methods. Therefore, the proposed method based on the effective area model is not only more realistic but can also help pipeline operators avoid cost wastage caused by excessive reliability management.



(a) Defect 1 (b) Defect 2

Figure 5.16 Reliability comparison of the traditional method and proposed method



(a) Defect 1 (b) Defect 2

Figure 5.17 CR comparison of the traditional method and proposed method

5.3.3 Configuration of RK4-ResNet

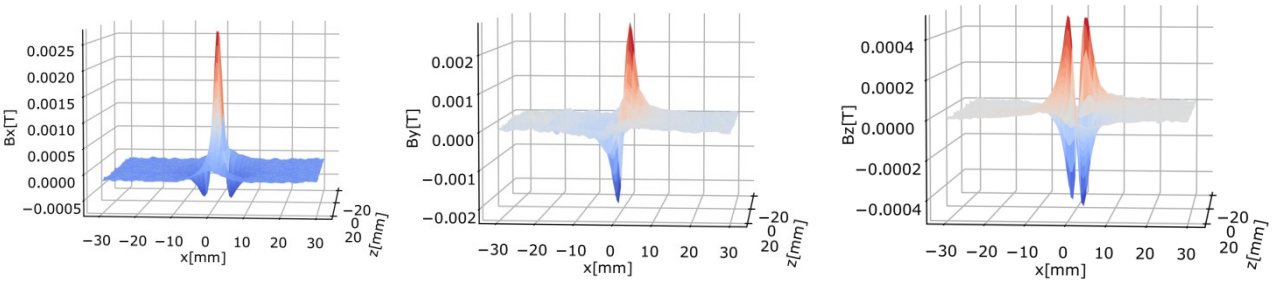
In this part, we explore the effect of various network architectures on the model's performance. We mainly focus on the effect of the number of residual blocks. The most accurate model, the RK4-ResNet model with one to three residual blocks, is investigated in this study, and testing is conducted using the same dataset. RK4-ResNet performs steadily better as the number of residual blocks rises, according to the test results shown in Table 5.3. However, the RK4-ResNet model with three residual blocks does not substantially outperform the RK4-ResNet model with two blocks in terms of prediction outcomes. Furthermore, compared to RK4-ResNet with two residual blocks, RK4-ResNet with three residual blocks requires greater processing time. Therefore, we can use RK4-ResNet with two residual blocks to perform the reliability analysis task. In addition, we also analyzed the model performance of Euler-Net with two residual blocks and CNN with the same depth as two residual blocks from the MSE perspective, and the conclusions reached are consistent with those in Section 5.3.1.

Table 5.3 Experimental results with different network architectures

Model	Max loss	Min loss	Avg loss	Training time(s)
RK4-Net (one block)	1.996e-1	1.387e-5	7.148e-2 ± 6.375e-2	2558.14
RK4-Net (two blocks)	9.868e-2	2.009e-8	2.957e-3 ± 2.887e-3	4634.20
RK4-Net (three blocks)	9.574e-2	2.158e-8	2.753e-3 ± 2.176e-3	6734.17
Euler-Net (two blocks)	1.985e-1	2.008e-8	5.7-9e-2 ± 5.522e-2	4420.28
CNN	4.356e2	2.116e-3	1.089e2 ± 1.065e2	4188.23

5.4 Industrial application and model validation

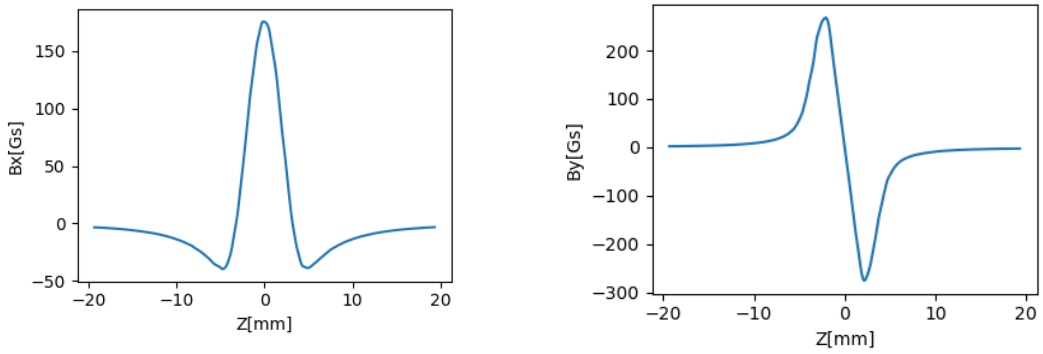
In order to minimize the interference of non-defective factors, the effect of the background magnetic field was eliminated, and the field data was filtered. Take the sample defect shown in Figure 5.5 as an example. Figure 5.18 demonstrates the MFL signal of the defect, which is consistent with the simulation data, illustrating the correctness of the FE model.



(a) Bx (b) By (c) Bz

Figure 5.18 The MFL signal of the defect

To further demonstrate the validity of the model, we compared our simulated results with the experimental results obtained in [300]. The parameters in the proposed FE model were adjusted to match those specified in the reference. The obtained results shown in Figure 5.19 closely align with the results presented in [300] for a lift-off distance of 1.10 mm, verifying the correctness of the establishment and calculation of the proposed model. Meanwhile, the results obtained from the proposed model are also consistent in terms of waveform characteristics when compared with other published articles [301-303].



(a) Bx (b) By

Figure 5.19 Simulation results for model verification

The MFL signal data shown in Figure 5.18 are put into the trained model for calculation, and the obtained results are shown in Figure 5.20. It can be found that the reliability predicted by the model built based on the ResNet family for the future 30 years is basically the same as the results obtained by MCS, which also indicates the proposed model's validity. Besides, the performance of the RK4-Net is better than the Euler-Net, which is consistent with the simulation data analysis.

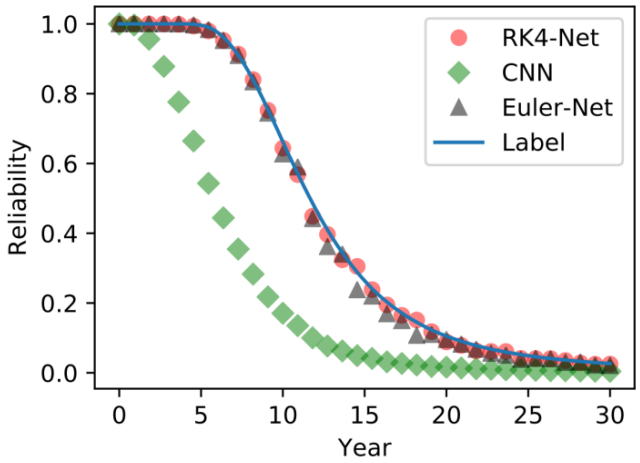


Figure 5.20 The obtained results from different methods

In engineering practice, the cost can be evaluated by the related experts. Therefore, according to Eq. (20)-(25), the relationship between CR and the pipeline's failure probability and the optimal point that minimizes the cost rate is shown in Figure 5.21. The corresponding failure probability at the optimal point is 2.49×10^{-3}. Hence, based on the reliability curve over time, we can derive the optimal re-assessment time interval of 5.33 years.

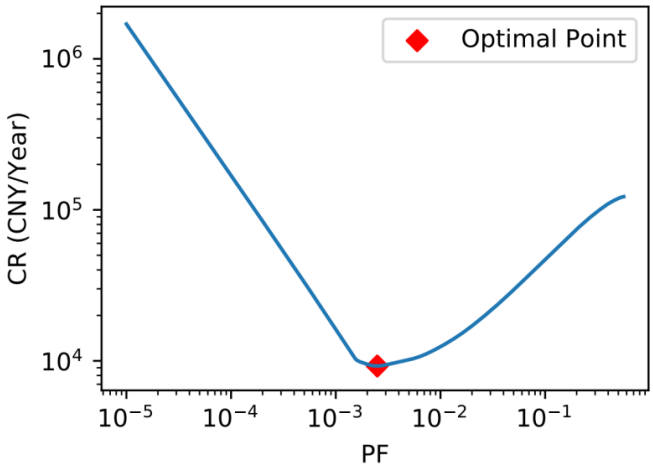


Figure 5.21 The relationship between CR and failure probability

5.5 Conclusions

This research presents a novel method for the reliability management of the corroded pipe. This chapter establishes a complete algorithm framework for mapping MFL signals to the pipeline's reliability, optimal re-assessment time interval, and CR, which

can improve the efficiency of risk and reliability analysis while ensuring the model's accuracy. Consequently, the results of the suggested method can be an effective pipeline integrity management tool that allows operators to evaluate the risk condition of corroded pipes after MFL inspections and promptly undertake the appropriate repair steps.

In this chapter, the defect's effective area is taken into consideration instead of the defect depth in the traditional methods when establishing the algorithm framework, leading to more accurate results. The ResNet-based data-driven method demonstrates its effectiveness for reliability analysis. The data set is simulated from an FE model, and an industrial application also verifies the model's validity. The three-axial MFL signals are used as the data-driven model's input. Through analysis, the performance of the neural networks with different architectures is in order: RK4-Net $\gtrsim$ Euler-Net $\gg$ CNN. The case study also provides guidance on how to apply the proposed method in practice. From the MFL inspection signal, the pipeline operators can directly obtain the reliability and their changing trends of the pipeline under inspection, as well as suggestions for optimizing costs and re-assessment intervals. The proposed model can be used to support the digitization management of pipeline operations.

The proposed method is a promising approach to reliability prediction of corroded pipes. However, the following are some limitations. (1) Only regular defects are currently considered, and irregular defects are not studied. (2) Only independent defects are considered without analyzing the interaction rules of multiple corrosion effects. (3) During the process of cost analysis, only the most significant aspects are taken into account, and some fixed expenditures are simply ignored. Therefore, further research needs to focus on addressing these concerns and including more actual detection samples in the training set to make more improvements. Besides, in this study, the methodological framework proposed for mapping MFL signals to reliability values, optimal re-assessment interval, and cost rate is generalizable. The specific set of parameters, including pipeline attributes and probabilistic inputs in the FE model and reliability analysis, are used as an example to illustrate the feasibility of the proposed

method. Furthermore, for different engineering application scenarios, these parameters can be modified to re-train the ResNet and obtain different results. Also, in future research endeavors, on the basis of the proposed model, transfer learning is a potential technique to handle more diverse tasks and obtain the desired results more flexibly.

Chapter 6: Reconstruction of 3-D pipeline defect profile based on MFL signals and hybrid neural networks

6.1 Introduction

The structural integrity of pipelines gradually deteriorates over time due to various loading conditions or environmental variables. This may result in incidents like leaks, which can cause personal injuries, significant property damage, and contamination of the environment [304, 305]. As a result, it is essential for pipeline operators to monitor the pipeline's conditions and carry out periodic inspections of the pipeline using a variety of inspection methods in order to diagnose damage and guarantee the pipeline's safety. The most commonly used pipeline inspection technique is non-destructive testing (NDT), which is also one of the essential steps in preventing pipeline incidents [306].

NDT techniques, including radiographic testing (RT), UT, and MFL, are widely utilized to identify pipeline defects [307]. The MFL procedure, compared to other inspection methods, is simple, has deficient requirements for the testing environment, and can detect a wide range of faults. It is now the most extensively used pipeline fault diagnostic technology [308]. The primary principle behind MFL is to generate a detection signal source by magnetizing a target magnetic conductor, such as a pipe, and then identify the defects by collecting the leakage flux beyond the pipe wall using Hall sensors [309]. When the ILI tool moves over a defect, the Hall sensor can detect an abnormal leakage flux. Defect size, shape, and other information can be determined by analyzing the characteristics of the MFL signal.

One of the most critical objectives in processing the MFL information is reconstructing the defect's three-dimensional (3-D) profile. The defect geometry may

be used to direct industrial maintenance and integrity management. Therefore, accurately estimating the defect's 3-D profile is of great significance for the safe and efficient operation of pipelines.

Many scholars have conducted related studies, as summarized in Table 6.1, which can be classified into four stages: defect identification, defect parameter estimation, 2-D defect profile reconstruction, and defect 3-D profile reconstruction. In Table 6.1, [7-11] focus on simply classifying defects and non-defects using deep learning methods but cannot quantify the defects precisely. Moreover, [12-18] use FE methods, machine learning methods, and signal processing methods to estimate and recognize the size of defects, [19-22] use machine learning and geometric methods for the 2-D and 3-D reconstruction of defects.

The following can be found from the research presented above. (1) Most recent research on MFL signals focuses on problems with classification and the quantification of defect sizes; relatively few of these studies include 3-D reconstructions of defect profiles. The defect size quantification only includes estimating the defect length, width, and depth rather than determining defect shapes, such as cuboid, spheroid, etc., thus reducing the accuracy and comprehensiveness of the results. (2) Some existing studies limit themselves to analyzing the single-axis MFL signals or use only the axial and radial signals as the basis for analysis. In fact, the MFL signals in the circumferential direction also contain more detailed features that are not sufficiently characterized by other directions and should not be neglected. The analysis based on the three-axes MFL signal is beneficial for obtaining more comprehensive information about the defect. The extraction of only single-axis signals has information deficiencies, which affects the accurate estimation of defects. (3) Regarding the data sources, a few studies have used magnetic dipole models and retrieved the data via simulations. The model's simplicity and ease of calculation are evident benefits, but an obvious limitation is its insufficient simulation accuracy, making it difficult to achieve correct reconstruction results. Some studies use field data to perform the analysis. The small number of defect samples in some actual inspections and the insufficient number of defect samples obtained in the

pull test may also lead to inaccurate results during the modeling. (4) The shapes that can be reconstructed by the existing defect 3-D profile reconstruction methods are limited, mainly consisting of basic shapes such as hemispheres and cylinders, where the defects are equal in length and width. Moreover, in actual situations, the defects may not align perfectly with either the horizontal or vertical plane, thereby resulting in a certain degree of rotational displacement. Currently, there is a lack of techniques available for reconstructing defects with rotation angles.

Table 6.1 Some relevant studies for pipeline defect identification based on MFL

Literature	Proposed method	Problems solved
[310]	Modified fuzzy min–max neural network	Classify the normal and abnormal conditions in the pipeline
[311]	Sparse self-coding, CNN	Classify of girth welds and spiral welds
[312]	CNN with the rectified linear units (ReLUs) employed as the activation functions	Classify the response segments into defect, cathodic protection or tee
[266]	Multiscale shot multibox detector network	Automatically identify the location of girth weld, spiral weld, and defect
[79]	CNN with normalization layer added	Identify injurious or noninjurious defect
[86]	Pattern-adapted wavelets and ANN	Estimate the length and depth of metal-loss defects
[267]	Improved Particle Swarm Optimization and RBFNN	Recognize the width and depth of defect
[313]	Nonlinear 3-D finite-element method	Recognize the depth of defect
[85]	ANN	Estimate the depth of defect
[314]	Physics-informed doubly fed cross-residual network	Quantification of defect length, width, and depth
[80]	Visual Transformation Convolutional Neural Network	Estimate the defect size
[315]	Heterogeneous Multiclass Feature Fusion	Estimate defect length, width, and depth
[300]	Characteristic approximation approach	Recognize the opening profile
[303]	Visual deep transfer learning neural network	Estimate defect size and cross-sectional profile
[301]	Geometric parameters based rational Bézier curve model	Fast reconstruction of 3-D defect profile
[82]	Iterative neural network	3-D defect profile inversion

This chapter suggests an approach for reconstructing the 3-D pipeline defect profile using MFL inspection signals. Compared to previous studies, the contributions

are summarized as follows: (1) A novel hybrid neural network-based 3-D defect reconstruction method is proposed, which can directly inverse the defect shape. Compared with the traditional methods, the model's accuracy is significantly improved. (2) The three-axis MFL signals are used as model input. The neural ODE maps the MFL signals to the spatial position of each point on the defective concave surface based on its excellence in parameterizing a homeomorphism between two sets. Furthermore, the model incorporates the Fourier integration kernel to enhance computational efficiency. (3) Due to the difficulty in obtaining sufficient amounts of high-quality experimental data, the proposed model can be trained on data from FE simulations and then transferred to the experimental dataset, effectively solving the problem. (4) The proposed method can well reconstruct not only the defects but also their rotation angles.

The rest of this work is structured as follows: Section 2 introduces the principle of MFL inspection; Section 3 presents the FE model establishment process, the basic concepts, and the proposed method; Section 4 discusses the results, model performance, and robustness; Section 5 provides an experimental application to illustrate the proposed method further; and Section 6 presents the conclusion.

6.2 Principle of MFL inspection

MFL inspection has been widely recognized as a proficient method for inspecting pipelines composed of ferromagnetic materials. It enables the identification, localization, and quantification of pipeline defects caused by metal loss [316]. The principle is based on magnetism. When two magnets with opposing polarities of significant strength are put in proximity to the pipe wall, the pipe becomes magnetized. This results in magnetic flux lines through the pipe wall, moving from the south pole toward the north pole. In the event of a metal loss defect in the pipe wall, a new north pole and a new south pole can be observed at the edges of the metal loss area [317]. This phenomenon is akin to the behavior of a magnet that has been fractured into two separate pieces. This occurrence can be attributed to various factors, such as corrosion. Consequently, the magnetic flux lines traverse the pipe wall, passing through the defect-

created air gap and then again through the pipe wall, as illustrated in Figure 6.1.

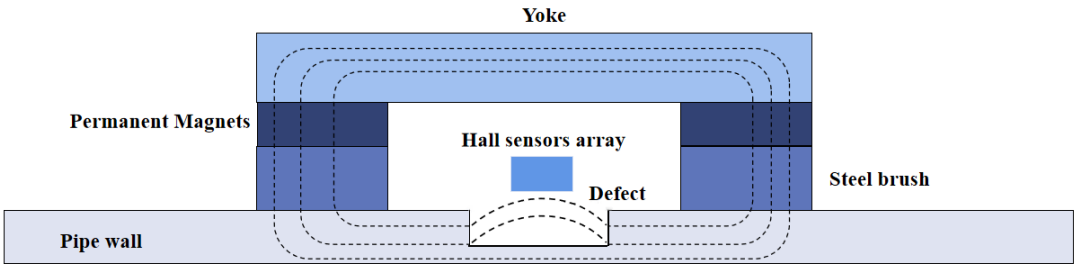


Figure 6.1 The diagram of MFL inspection principle

The magnetic flux density within a given material is contingent upon its magnetic permeability, a fundamental physical property. Various materials exhibit varying degrees of magnetic permeability. It can be observed that ferromagnetic steel exhibits a significantly greater magnetic permeability in comparison to air. Consequently, the magnetic flux density in the pipeline's sections without defects is considerably greater than in the air gap region resulting from the defect. Therefore, the magnetic flux lines will protrude while traversing the air gap due to the limited flux capacity of air per unit volume. This phenomenon is called magnetic flux leakage (MFL), which pipeline operators commonly use to conduct in-line inspections [302].

The MFL-ILI system primarily comprises permanent magnets, Hall sensors, steel brushes, and a yoke. Through the steel brushes, permanent magnets magnetize the pipe wall to the saturation stage. The sensors are positioned uniformly around the pipe's circumference and are propelled by the ILI tool in the pipe's axial direction. The sensors capture MFL signals, which are subsequently recorded and subjected to analysis for the purpose of identifying potential anomalies. Different defect shapes or sizes will generate leakage magnetic fields with varying distribution characteristics. Hence, defect reconstruction can be achieved by analyzing the recorded MFL signals. Furthermore, the yoke serves to enhance the magnetic circuit architecture and diminish the ambient magnetic field while conducting inspection.

It is common practice to use an approximation of the leaking magnetic field around the defect for research. Typically, modeling complicated geometric defects using the FE simulation may provide more precise results to resolve the magnetic field problem.

The region under study is partitioned into numerous discrete segments, enabling the computer to calculate Maxwell's equations at each point.

6.3 Proposed method

The overall model architecture and specific parameters are shown in Figure 6.2 and Table 6.2. The model architecture proposed in this study consists of several modules:

- The input and output module
- The feature extraction module utilizing the Dual-scale CNN algorithm
- The depth prediction module
- The canvas size prediction module

The canvas represents the minimal surface area within the X-Z plane capable of displaying defects in any orientation, as shown in Figure 6.3. The depth prediction module is further divided into the mask and Neural ODE modules. The model's inputs are the MFL signals obtained in three directions, which can be considered 100*100 pixel images. By integrating the depth of each point within the canvas and the size of the canvas, as shown in Figure 6.3, we can finally obtain the defect's 3-D profile. Notably, this approach enables the acquisition of 3-D profiles of defects at arbitrary angles rather than solely parallel to a specific direction.

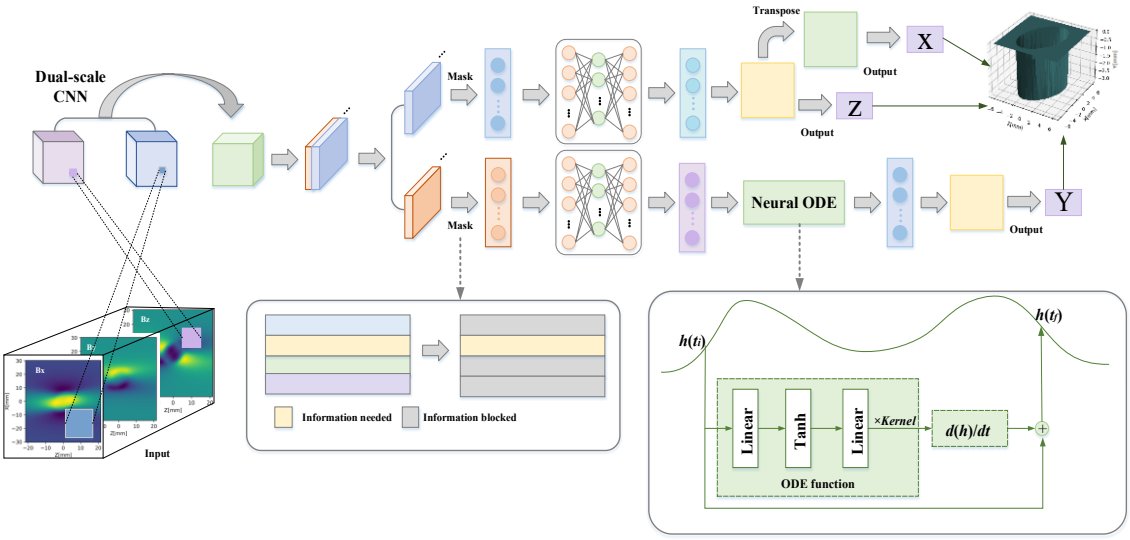


Figure 6.2 The architecture of proposed model

Table 6.2 The network parameters of proposed model

Critical layers	Type	Output shape
Input layer	-	3*100*100
Feature extraction module	Conv2D (kernel size 5*5, 3*3)	64*100*100
	Mask	1*80000
Depth prediction module	Neural ODE	1*80000
	Dense	100*100
	Mask	1*80000
Canvas size prediction module	Dense	100*100
	Transpose	100*100
Output layer	-	3*100*100

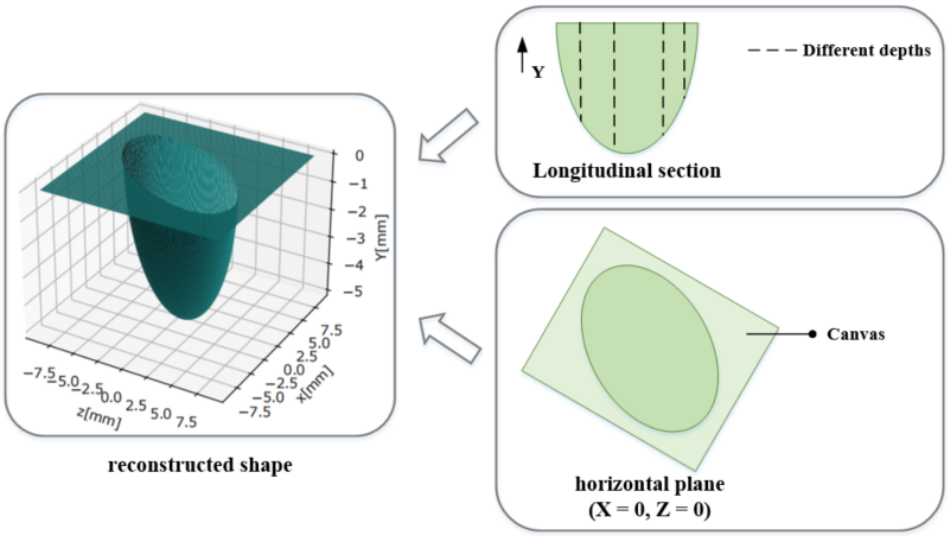


Figure 6.3 Image interpretation of defect reconstruction

First, based on FEM, we designed four different shapes of defects. Furthermore, for each shape, we considered various defect sizes. The MFL signals of each defect in three directions are acquired through simulation and utilized as inputs to the model. The dual-scale convolution is employed in the feature extraction module to enhance the network model's capacity to comprehend complex data features. This is achieved by incorporating convolution kernels of varying sizes to capture feature information effectively. In contrast with traditional convolutional layers using single-scale convolutional kernels, the approach employed in this study provides boosted feature representations, thereby yielding accurate results.

The depth prediction module incorporates a masking mechanism that functions as a tensor filter. This mechanism effectively reduces the model's attention to irrelevant

data, resulting in a decrease in the number of parameters within the model. Consequently, the efficiency and performance of the model are enhanced. The Neural ODE algorithm is used to predict the defect depth associated with the points on the canvas. As depicted in Figure 6.3, with the exception of blocks and cylinders, the depth of each point within the canvas exhibits variation. Therefore, instead of predicting a single depth value, it is necessary to indicate the depth of every individual point, i.e., the value of each pixel, on the canvas in order to achieve the intended 3-D reconstruction.

Using traditional neural network techniques, including ANN, CNN, and ResNet, for defect 3-D reconstruction necessitates the training of a substantial number of parameters, which consumes a significant amount of graphics ram resources. However, parameter sharing in Neural ODE exhibits notable characteristics that effectively address this limitation. Furthermore, according to Maxwell's equations, the solution to the electromagnetic field is unique. Consequently, it is achievable to reconstruct the defects by establishing the correlation between magnetic field and spatial quantities. According to [318, 319], neural ODE is more appropriate for parameterizing a homeomorphism between two sets. Therefore, this chapter uses a neural ODE-based method for defect reconstruction. Still, the utilization of neural ODE may result in a notable reduction in the solution speed as the model complexity increases during the training phase. To address this issue, this study incorporates a Fourier integral kernel, which effectively enhances the ODE solver's solution time stability.

6.3.1 FE model establishment

According to the principle of MFL inspection, the FE model is built as illustrated in Figure 6.4. The parameters related to the FE model is the same as in Section 5.2.2.1. The size of the defect in the pipe's axial, circumferential, and radial directions is defined as the defect's length, width, and depth, respectively. Figure 6.4 shows the 3D schematic illustration of the MFL signal of a cuboid defect. In addition, we also simulated defects in cylindrical, ellipsoidal, and truncated cone shapes with different orientations. Among

them, some standard shapes are also included. For example, ellipsoids and truncated cones has bodies with cross-sections of ellipses and circles, and truncated cones also contain cones. Take cylindrical defects as an example. Here, we use heatmap plots to visualize the MFL signal characteristics, as shown in Figure 6.5, which demonstrate the simulated MFL signals for cylindrical defects in different directions. Meanwhile, we mapped the contour projection of the defect in the y-direction on heatmap plots.

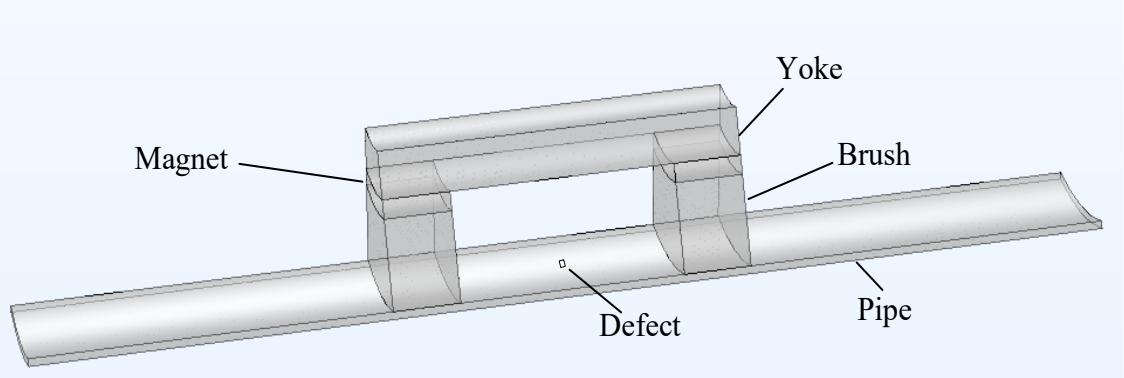
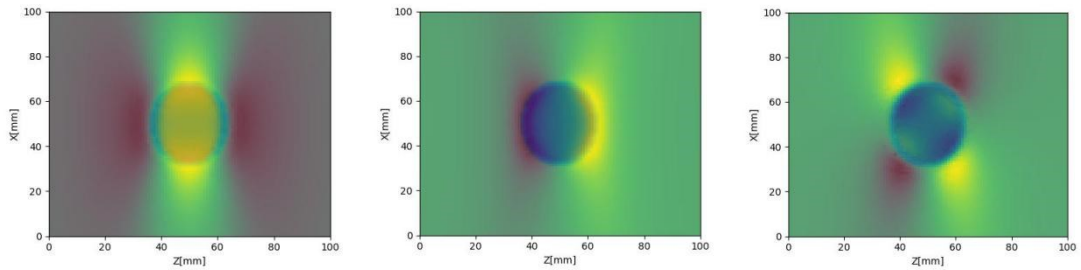


Figure 6.4 The FE model



(a) Bx (b) By (c) Bz

Figure 6.5 Heatmap plots of MFL signals of cylindrical defect

A total of 4842 data sets are obtained by using the technique of FE simulation. Of the total samples, 90% are assigned at random for training the algorithm, while the remaining 10% are set apart for validating the effectiveness of the model.

6.3.2 Dual-scale CNN module

Dual-scale refers to extracting features at two different scales, thereby enabling the capture of different levels of information within the image [320]. Additionally, the network's robustness is enhanced through parameter sharing, thus mitigating the

potential for overfitting.

In this chapter, the images obtained from the FE simulation are initially fed into the convolution layers, consisting of convolution kernels of varying sizes. This chapter utilizes two convolution kernels, one with a size of 3 × 3 and the other with a size of 5 × 5. These kernels are employed to extract feature maps, denoted as X and Y, respectively. Due to the larger kernel size, feature map Y possesses a greater receptive field. Consequently, feature maps X and Y are combined to yield feature map Z, as shown in Eq. (6.1).

$$Z = X \oplus Y \tag{6.1}$$

where the symbol $\oplus$ denotes element-wise summation.

The dimensionality of feature map Z is equivalent to that of X and Y. However, due to the mechanism of dual-scale CNN, each dimension of feature map Z encompasses a greater amount of information.

6.3.3 Mask module

This section focuses on the mask module. "Mask" is prevalent in various applications, including capsule networks and Transformer models [321, 322]. The utilization of a "mask" is employed to limit the operations of a model or modify the allocation of attention within the model to cater to the specific requirements of a given task. By applying an appropriate mask, certain restrictions can be imposed on the model outputs, thus affecting their delivery and use in the network. This chapter's data inputs are MFL signals of four defect shapes. Hence, the mask mainly serves as a sensor filter. This filter selectively masks or retains the original tensor, directing the model's attention toward the shape with the highest category affiliation so as to facilitate deeper learning of defect depths within the specified category. The "mask" can enhance the network's capacity to effectively acquire and analyze crucial data, thereby improving the interpretability of the neural network.

6.3.4 Neural ODE module

Neural Ordinary Differential Equation (Neural ODE) is a deep learning method based on Ordinary Differential Equation (ODE), combining traditional numerical solution techniques and neural network models [323]. ODE is used to describe the relationship of an unknown function with the change of variables. Neural ODE establishes a connection between neural networks and the initial value problem (IVP) associated with ODEs. The input state in Neural ODE is the initial value of the ODE, and the neural network can be regarded as the driving function. The depth-solving process can be considered a continuous iterative optimization of the driving function, ultimately leading to the desired results [324, 325].

In this chapter, Neural ODE is used to predict the depth at different points within the canvas plane. Due to the prediction's complexity, many parameters are required when using traditional methods for prediction, resulting in a heavy computational burden. In contrast, Neural ODE has a simple structure with fewer training parameters, improving computation efficiency.

In traditional neural networks, such as ResNet, the underlying units can be represented by a generalized formula, as shown in Eq. (6.2) [326].

$$h_{t+1} = h_t + f(h_t, \theta_t) \tag{6.2}$$

where $h_t \in R^D$, t denotes the number of the discrete layer, $f(\cdot)$ is a differentiable function and θ means the learnable parameters of $f(\cdot)$. The basic expression of the Neural ODE is obtained by transforming the discretized form into a continuous form, as shown in Eq. (6.3) [327].

$$\frac{dh(t)}{dt} = f(h(t), t, \theta) \tag{6.3}$$

with initial value $h(t_0) = h_0$, h represents the state, f is the driving function, representing trainable layers parameterized by the weight θ, t falls within $[t_i, t_f]$, representing the virtual evolution time. The process of forward propagation for Neural ODEs can be viewed as the evolution of ODEs with the initial state over time.

Furthermore, considering the rapid increase in the number of iterations, the neural

network's fitted function will progressively exhibit greater complexity, resulting in an increased level of problem stiffness. Consequently, the solution's efficiency will be significantly diminished. Hence, this chapter introduces the Fourier integral kernel as a means to alleviate the problem's stiffness, thereby stabilizing the solution time. The specific expression is as follows:

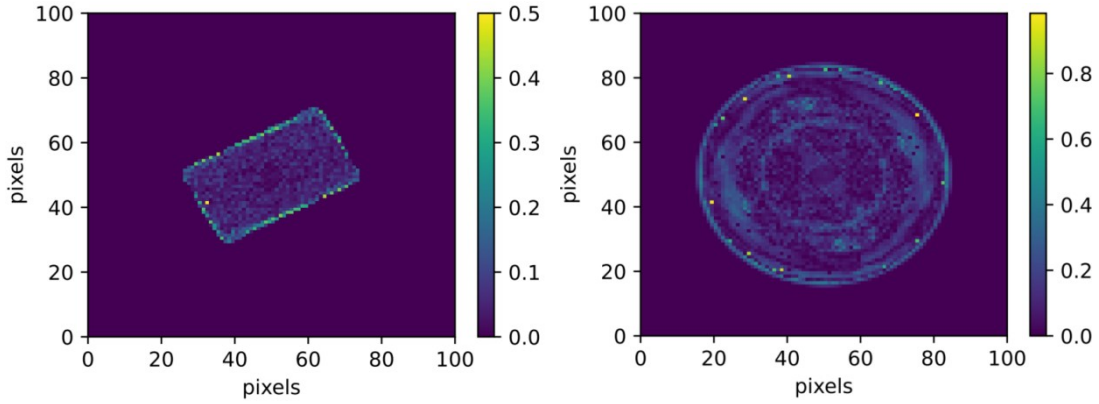
$$h(t_f) = \int_{t_i}^{t_f} f(h(t), t, \theta) sin(2\pi t) dt \tag{6.4}$$

6.4 Results and discussions

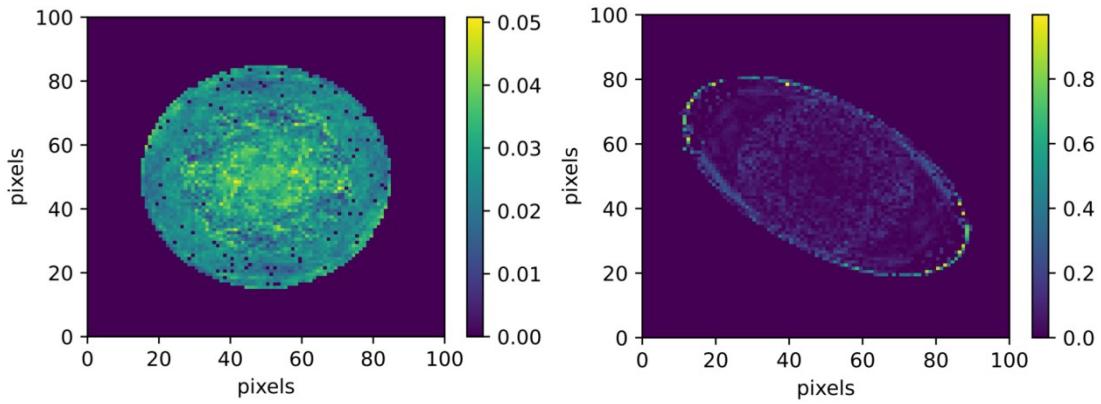
6.4.1 Results analysis

We randomly selected one defect from each shape for error visualization, as shown in Figure 6.6. Each grid corresponds to a pixel, with each pixel denoting the absolute value of the labeled depth at that point subtracted from the reconstructed depth. The more significant the difference in color between the pixel point and the canvas, the higher the inaccuracy.

In Figure 6.6, it is evident that all shapes exhibit slight variations in their overall measurements. With the exception of the cylinders, the other profiles demonstrate more significant errors in proximity to the edges of the defects. Hence, it is reasonable to infer that such errors might be attributed to the rotation angle.



(a) Cuboid (b) Truncated cone



(c) Cylindrical (d) Ellipsoidal

Figure 6.6 Error maps of randomly selected defects

6.4.2 Model performance comparison

This chapter uses mean absolute error (MAE) to estimate the accuracy of the proposed model. As shown in Table 6.3, we compared the proposed method with ResNet. The results illustrate that the accuracy of the proposed method is higher than ResNet on cuboid defects, cylindrical defects, ellipsoidal defects, truncated cone defects, and all the defects with mixed shapes.

Table 6.3 Comparison of different methods on reconstruction accuracy

Type	Proposed method	ResNet
Cuboid	0.0721±0.0253	0.5575±0.4329
Cylindrical	0.0293±0.0203	0.7937±0.4867
Ellipsoidal	0.0735±0.0267	0.2398±0.3015
Truncated cone	0.1281±0.0528	0.8369±0.5132
Mixed shape	0.0757±0.0486	0.6070±0.5009

Similarly, one defect is randomly selected from each of the four shapes for visualization purposes. The real profile of the defect, the profile reconstructed using the proposed method, and the profile reconstructed using ResNet are depicted in Figure 6.7. We can conclude the same as above that the proposed method exhibits a higher level of accuracy. Additionally, it can be observed that among the four different defect shapes, the reconstruction of cylindrical defects yields the most desirable outcomes.

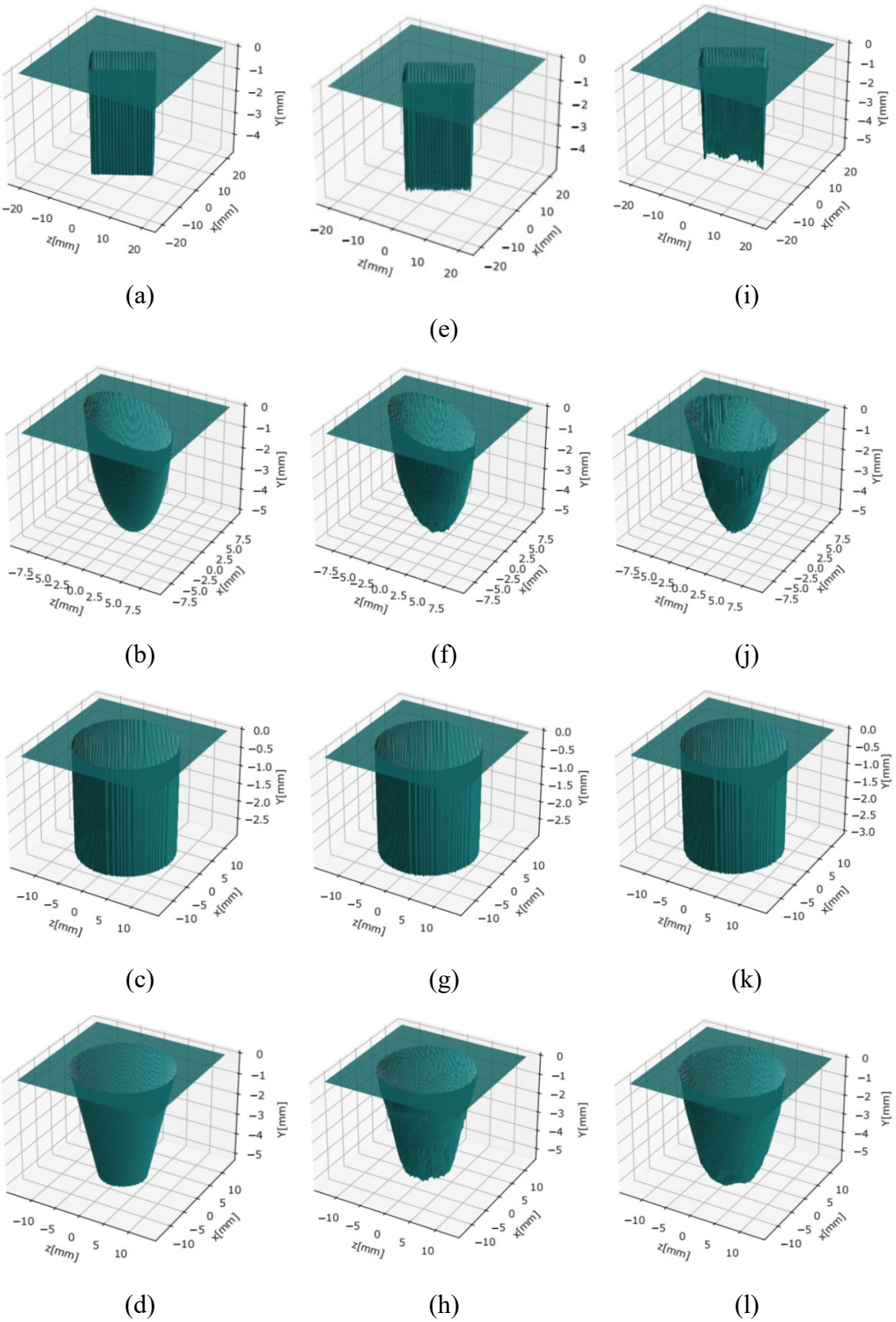


Figure 6.7 Comparison of different methods. (a), (b), (c) and (d) are the true defect profiles. (e), (f), (g) and (h) are the reconstructed defect profiles using proposed method. (i), (j), (k) and (l) are the reconstructed defect profiles using ResNet.

Furthermore, two methods, one using the Fourier integral kernel (the proposed method) and the original neural ODE method, are used for comparison. Based on 12 GB NVIDIA TITAN XP GPU, we ensure that the neural network is otherwise structurally consistent, comparing the time required to iterate between the two methods. As shown in Figure 6.8, the Y axis denotes the relative time, which is the ratio of the time taken by the original neural ODE to the time taken by the proposed method. It can be observed that with an increasing number of iterations, the relative time exhibits a progressive rise, indicating a significant decline in the solution efficiency of the original method compared to the proposed method. Therefore, the proposed method may substantially enhance the stability and efficiency of the solution process.

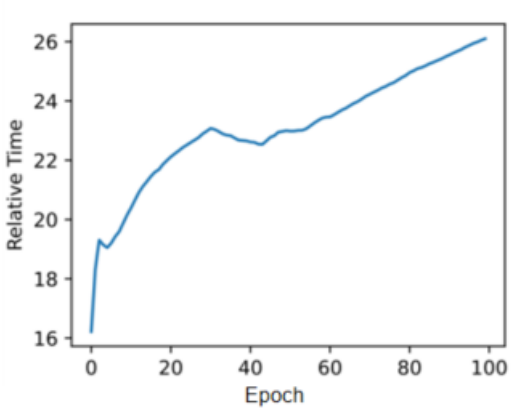


Figure 6.8 Relative time between original and proposed methods

6.4.3 Model Robustness Comparison

In practical applications, it is common for the obtained signals to be affected by noise owing to the complex conditions of the pipeline. To assess the robustness of the proposed method, we introduce Gaussian white noise with five levels into both the training and test samples. These noise levels are as follows: 1% (signal-to-noise ratio (SNR) = 20 dB), 3% (SNR = 15.23 dB), 5% (SNR = 13.01 dB), 7% (SNR = 11.55 dB), and 9% (SNR = 10.46 dB), respectively. The results obtained are shown in Figure 6.9. It can be observed that the proposed method demonstrates greater estimate accuracy compared to ResNet across various noise levels. It can also be noticed that the error

increases as the percentage of noise increases. However, for the proposed method, even when the noise percentage increases to 9%, the MAE is still not exceeding 0.2 mm.

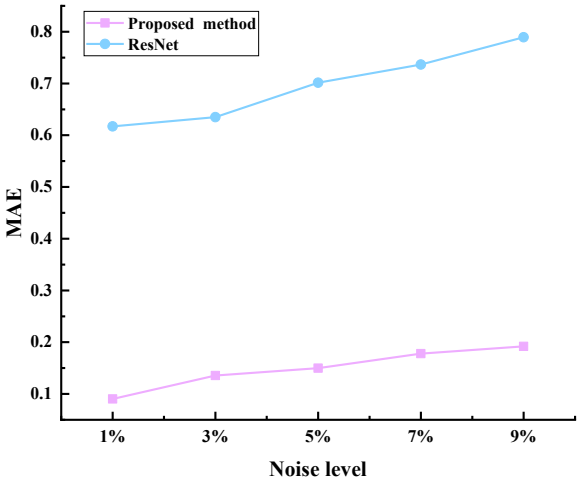


Figure 6.9 Robustness tests

6.5 Experimental analysis

In order to provide further validation for the composite neural network model suggested in this research, a field study is undertaken. Figure 6.10 displays the structure of the MFL tool used during the pull test. Fifty artificial defects with different geometrical parameters and shapes are machined on the outer surface of the pipe wall. The pipe's outer diameter is 323 mm, while its wall thickness is 8 mm. The dimensions of the defects vary between 5 mm and 25 mm in length and width and between 1 mm and 5 mm in depth. The detailed parameters of length, width, and depth are shown in Figure 6.11. The MFL tool is used to conduct a pull test and record the three-axis MFL signal produced by the manufactured defects. To reduce the influence of non-defective factors, the impact of the background magnetic field is filtered.

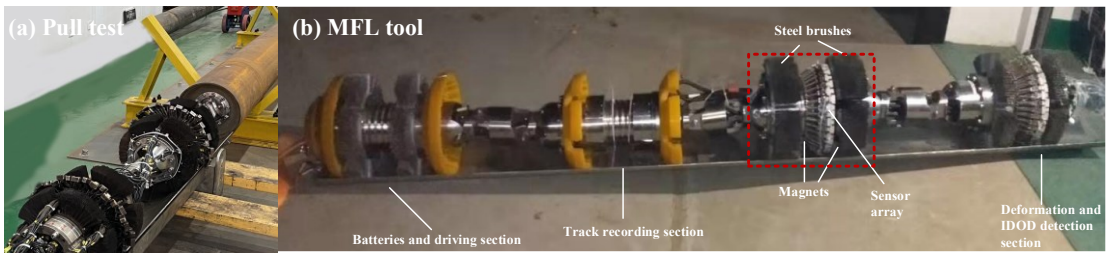


Figure 6.10 The experimental situation

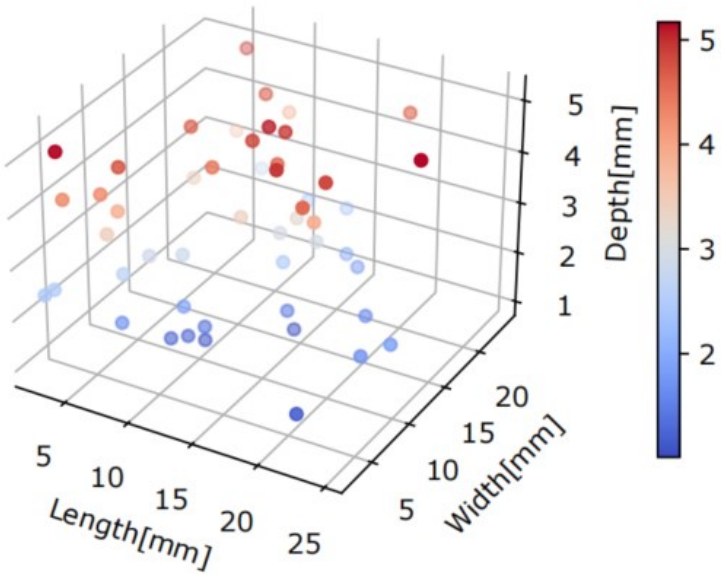


Figure 6.11 Defect parameters' distribution

The trained model derived from the simulated data of FE analysis is regarded as a pre-trained model. This pre-trained model is then transferred to the dataset obtained from the pull test for practical implementation. Out of the total datasets, 40 are allocated for training, while the remaining 10 are designated for testing. Given that the pre-trained model has already acquired a general feature representation of the defect profile, the quantity of data needed for fine-tuning the pull test dataset is often lower compared to training the model from the beginning. This phenomenon is especially advantageous in the context of industries characterized by limited availability of data.

The errors acquired by the suggested method and ResNet are compared, as demonstrated in Figure 6.12. The average defect reconstruction error of the proposed method is 0.1018 mm, indicating a lesser value compared to that of ResNet. These results align with the conclusions drawn in the previous section. Figure 6.13 depicts the actual image of the defect, the actual contour of the defect, and the reconstructed profile of the defect. It can be found that the reconstructed defect profile can accurately characterize the real defect profile. Therefore, both simulation and experimental studies show that the proposed method can accurately reconstruct 3-D defect profiles from three-axis MFL data.

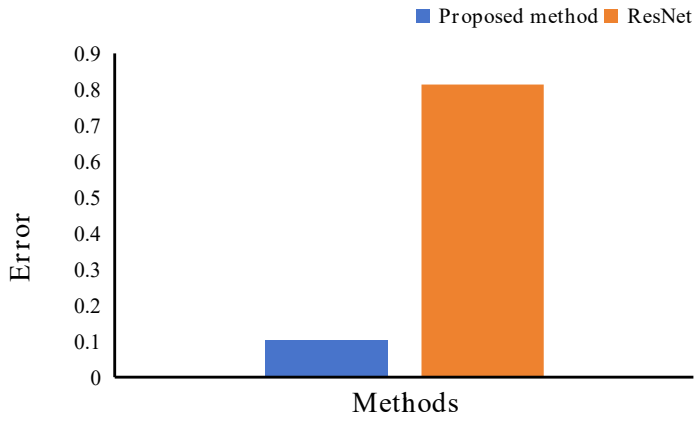
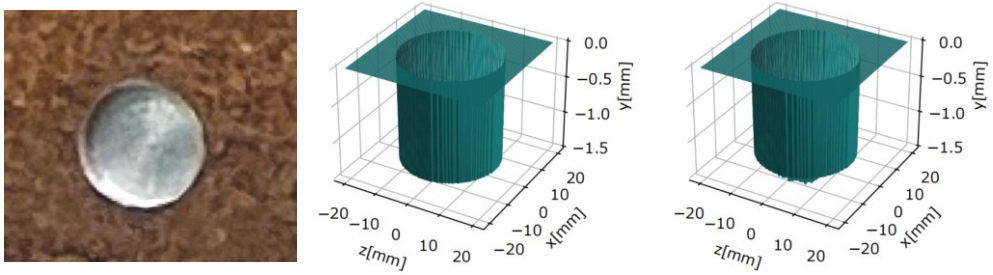


Figure 6.12 Accuracy comparison of the proposed method and ResNet



(a) (b) (c)

Figure 6.13 Comparison of defect profiles. (a) The actual image of the defect. (b) The actual profile of the defect (c) The reconstructed profile of the defect using the proposed method.

6.6 Conclusions

This study proposes an innovative approach for reconstructing the defect profile of oil and gas pipelines by using MFL data. It presents a novel hybrid neural network architecture for accurately and efficiently mapping three-axial MFL signals to the 3-D profile of defects. The proposed method is compared with existing methods to demonstrate its superior performance. As a consequence, the outcomes of the proposed method may serve as an effective tool for intelligent integrity management in pipeline systems. This tool enables the visualization of defects in underground pipes via the analysis of MFL signals, hence facilitating operators in undertaking subsequent maintenance measures.

In this chapter, we design four different shapes of defects with various sizes using

FE simulation and obtain the MFL signals of each defect in three directions as inputs to the model. In the construction of several modules for the hybrid neural network, we use a two-scale convolution module, a masking mechanism, and a neural ODE with a Fourier integral kernel. The final 3-D profile of the defect is created by integrating the predicted canvas size and the depth measurements of every point on the canvas. Each module has specific advantages and contributes to the feature extraction process, enhancing the model's computational efficiency, conserving graphics ram resources, and introducing novel ideas for defect reconstruction. The model trained on FE simulation data is successfully transferred to the pull test dataset with desirable outcomes. The simulation results indicate that the proposed method is able to reconstruct defect profiles with an average error of 0.0757 mm, while the pull test results show that the defect reconstruction average is around 0.1mm. Furthermore, the proposed model does not impose any limitations on the orientation of the reconstructed defects, allowing for the reconstruction in any possible direction. The proposed model can be used to support the digitization management of pipeline operations.

The hybrid neural network suggested in this study has great potential for reconstructing defects' three-dimensional profiles. However, there are limitations to consider. (1) Presently, the focus of research is primarily on regular defects of a certain shape, whereas irregular defects remain unexplored. (2) The analysis just focuses on individual defects without considering the interaction rules of numerous corrosion effects. Hence, it is essential for future studies to prioritize solving these challenges in order to promote further advancements.

Chapter 7: Conclusions and future work

7.1 Conclusions

The primary aim of conducting reliability analysis and risk analysis for OGP is to assure the long-term functionality of the OGP systems and enhance safety measures while also reducing possible risks and associated losses. This is also key to protecting the environment, maintaining equipment, reducing costs, and increasing public trust. This Ph.D. thesis provides an in-depth discussion on the reliability and risk of different pipeline systems to address the challenges currently in the pipeline risk field, including the inappropriateness of the existing models with some OGP systems, complex model structures, low computational efficiency, and inaccurate analysis results. This thesis adopts systematic research approaches, integrates multiple data sources and technical tools, combines deep learning algorithms and traditional risk analysis methods, proposes risk and reliability analysis models for the characteristics of different pipeline systems from a data-driven point of view, and validates them with actual cases. This thesis not only proposes new OGP risk and reliability analysis frameworks but also extends the traditional analysis theory. The new theoretical models fill the gaps in existing research and provide novel ideas and paradigms for the reliability and risk research field of OGP. The research results of the paper contribute to the progress of scientific theories and provide guidance and decision support for practical engineering applications. The results can help pipeline operators to ensure the safe and smooth operation of OGP systems, discover the potential risks of the system in time, reduce the probability of accidents, optimize the cost and resource allocation, and establish the foundation for the digital and intelligent management of OGP.

In this thesis, different types of data related to OGP operations are used to build more reasonable risk and reliability analysis models for different pipeline systems: pipelines where ILI can be conducted, pipelines where ILI can not be conducted, and

transmission station.

The main contributions of each topic are summarized as follows:

Topic 1: Risk assessment of buried gas pipelines based on improved cloud-variable weight theory

- A novel cloud-variable weight model is proposed to conduct pipeline risk assessment research. Compared with the constant weight theory mentioned above, the variable weight theory (VWT) can highlight the negative effects of the index with higher risk, which can obtain more reasonable results.
- CM is used to improve the calculation in VWT, and it is applied throughout the entire risk evaluation process. The uncertainty can also be better conveyed through the calculation between CMs, making the expression of uncertainty more adequate and accurate.
- The results are significant to help pipeline operators identify the key pipe sections and risk factors in the system, optimize the resource allocation and improve the efficiency of risk management.

Topic 2: Discovery of potential risks for the gas transmission station using monitoring data and the OOBN method

- A structure mapping method based on FMEA is proposed. An OOBN framework is developed based on the overall system's process flow, making the model more simplified and flexible.
- Both leaky noisy-or gate (LNG) and expectation maximum (EM) algorithm are incorporated into BN's parameter learning, efficiently dealing with imperfect historical records and limited sample data while reflecting the uncertainty of risk factors.
- An accident precursor identification approach based on PAA-CUSUM is presented to identify possible vulnerabilities in the system in real-time by finding abnormalities using monitoring data. This allows backward analysis when anomalies are detected, which is more consistent with actual conditions than assumption-based analyses.

Topic 3: Reliability analysis of corroded pipes using MFL signals and Residual Neural Networks

- The FE simulation of MFL signals and reliability analysis are combined through deep learning methods, allowing the MFL signals to be directly mapped to the pipeline's reliability. The principles of reliability calculation based on the effective area model rather than just depth are also effectively integrated into this mapping process, which can obtain more accurate results.
- A novel ResNet-based reliability prediction method is proposed, making the analysis process more accurate and efficient. The case in this thesis shows that, compared to traditional methods, the proposed model's accuracy is more than 20% higher, and the computational efficiency has been increased by 200 times.
- The re-assessment interval optimization method is also integrated into the deep learning algorithm, which can automatically generate the optimal time interval and can effectively help pipeline operators make decisions based on the MFL inspection signals.

Topic 4: Reconstruction of 3-D pipeline defect profile based on MFL signals and hybrid neural networks

- A novel hybrid neural network-based 3-D defect reconstruction method is proposed, which can directly inverse the defect shape. Compared with the traditional methods, the model's accuracy is significantly improved.
- The three-axis MFL signals are used as model input. The neural ODE maps the MFL signals to the spatial position of each point on the defective concave surface based on its excellence in parameterizing a homeomorphism between two sets. Furthermore, the model incorporates the Fourier integration kernel to enhance computational efficiency.
- Due to the difficulty in obtaining sufficient amounts of high-quality experimental data, the proposed model can be trained on data from FE simulations and then transferred to the experimental dataset, effectively solving the problem. Besides, The proposed method can well reconstruct not only the defects but also their

rotation angles.

Regarding practical applications, the model from topic 1 is a valuable tool for assessing risks in urban underground gas pipeline systems. By establishing a comprehensive evaluation framework that connects various risk factors, regular consultations with experienced experts can provide periodic insights into pipeline risk status and maintenance recommendations. This approach can be extended to systems lacking precise data, such as defect information, to reflect the system's risk status. The results of the proposed model, which are the ranking of pipe sections' risk levels can be considered a prioritization when performing risk rehabilitation on this pipeline. Higher risk pipe sections should be prioritized for targeted control measures, depending on the key risk factors identified during the evaluation process. When two pipe sections have the same Ex value, the value of En should be evaluated.

The model proposed in topic 2 is well-suited for the GTS. By installing sensors to monitor parameters like flow and pressure, it can promptly identify anomalies that might lead to accidents, such as leaks or blockages. This model's adaptability extends to other complex systems with multiple components capable of obtaining monitoring data. The analysis results can provide the component's ranking of potential risk from high to low, as well as their possible failure types and failure causes, which can guide the GTS operator to perform efficient risk rehabilitation based on the failure types and causes of the components.

The deep learning models developed in topic 3 and topic 4 are suitable for oil and gas pipelines where ILI can be conducted. Collecting foundational data like pipeline wall thickness should be started first, and then MFL signals are derived through MFL tools. Using these trained deep learning models, operators can ascertain the pipeline's reliability, determine the optimal re-assessment period, and reconstruct 3-D profiles of defects. The advantage of deep learning models lies in their user-friendly nature, requiring nothing more than data input to yield insightful results. The obtained 3-D defect profiles can assist pipeline operators in gaining a more comprehensive

understanding of the internal conditions of the pipeline, enabling the formulation of more precise maintenance plans and providing support for establishing a digital twin system.

7.2 Limitations and future work

Despite the achievement of its goal, the research presented is confronted by particular limitations and can be addressed in future work:

- For the research of topic 1, the established risk assessment index system is only applicable to buried gas pipelines, but not to other types such as submarine pipelines or long-distance oil pipelines. Also, the proposed method is a static approach that only analyzes risk based on the current state of the pipeline and cannot dynamically track the development of risk.
- For research in topic 2, the risk factor classification can be more thorough. With comprehensive data, more in-depth analysis may be conducted in the future. Moreover, only the daily regular operation of the GTS system is considered, but not the risks under other operating conditions such as venting. Future studies on more complicated systems will require a more precise component separation, such as the process piping segregated according to pipe diameter and other factors. As a result, future studies should concentrate on these issues to ensure the stable and smooth operation of the GTS system.
- For research in topic 3 and topic 4, only regular defects are currently considered, and irregular defects are not studied, and only independent defects are considered without analyzing the interaction rules of multiple corrosion effects. In the future, more varieties of defects should be investigated. Besides, in topic 3, the methodological framework proposed for mapping MFL signals to reliability values, optimal re-assessment interval, and cost rate is generalizable. The specific set of parameters, including pipeline attributes and probabilistic inputs in the FE model and reliability analysis, are used as an example to illustrate the feasibility of the proposed method. Furthermore, for different engineering application scenarios,

these parameters can be modified to re-train the ResNet and obtain different results. In future research endeavors, on the basis of the proposed model, transfer learning is a potential technique to handle more diverse tasks and obtain the desired results more flexibly. For the 3-D reconstruction mentioned in topic 4, future efforts could explore the application of physics-informed neural networks. This involves incorporating the relevant physics laws into neural network models, allowing the neural network to learn the system's physical principles. The neural network structure is employed to approximate unknown physical scenarios, and the loss function considers constraints from the physics equations. By combining the flexibility of deep learning with the prior knowledge of physical equations, this approach aims to better address the challenges of complex defect reconstruction. Additionally, further investigation into alternative AI technologies is warranted. For instance, point cloud data generated from laser scans or other sensors can be utilized to detect surface defects in pipelines and related equipment. Deep learning models can then analyze and process the point cloud data to reconstruct the three-dimensional shape of objects. Moreover, exploring the potential of well-trained large-scale AI models designed for image processing and applying transfer learning methods may offer opportunities for defect 3-D reconstruction.

Bibliography

[1] Kraidi L, Shah R, Matipa W, Borthwick F. An investigation of mitigating the safety and security risks allied with oil and gas pipeline projects. Journal of Pipeline Science and Engineering. 2021;1:349-59.

[2] Canada NR. Pipelines Across Canada. https://www-nrcan-gc-ca.login.ezproxy.library.ualberta.ca/our-natural-resources/energy-sources-distribution/fossil-fuels/pipelines/pipelines-across-canada/188562020. (accessed May 29)

[3] Chen C, Li C, Reniers G, Yang F. Safety and security of oil and gas pipeline transportation: A systematic analysis of research trends and future needs using WoS. Journal of Cleaner Production. 2021;279:123583.

[4] Bonvicini S, Antonioni G, Cozzani V. Assessment of the risk related to environmental damage following major accidents in onshore pipelines. Journal of Loss Prevention in the Process Industries. 2018;56:505-16.

[5] Yin Y, Yang H, Duan P, Li L, Zio E, Liu C, et al. Improved quantitative risk assessment of a natural gas pipeline considering high-consequence areas. Reliability Engineering & System Safety. 2022;225:108583.

[6] Chen X, Wu Z, Chen W, Kang R, He X, Miao Y. Selection of key indicators for reputation loss in oil and gas pipeline failure event. Engineering Failure Analysis. 2019;99:69-84.

[7] Biezma MV, Andrés MA, Agudo D, Briz E. Most fatal oil & gas pipeline accidents through history: A lessons learned approach. Engineering Failure Analysis. 2020;110:104446.

[8] Coramik M, Ege Y. Discontinuity inspection in pipelines: A comparison review. Measurement. 2017;111:359-73.

[9] Brito AJ, de Almeida AT. Multi-attribute risk assessment for risk ranking of natural gas pipelines. Reliability Engineering & System Safety. 2009;94:187-98.

[10] European Gas pipeline Incident data Group, 11th Report of the European Gas pipeline Incident data Group [online], December 2020. Available from: https://www.egig.eu/reports (accessed 23 October 2021)

[11] Serious Incident Rate and Cause, PHMSA [online], 2020. Available from: https://www.phmsa.dot.gov/data-and-statistics/pipeline/national-pipeline-performance-measures (accessed 23 October 2021)

[12] Wu W, Yang C, Chang J, Chateau P, Chang Y. Risk assessment by integrating interpretive structural modeling and Bayesian network, case of offshore pipeline project. Reliability Engineering & System Safety. 2015;142:515-24.

[13] Zhang Y, Weng WG. Bayesian network model for buried gas pipeline failure analysis caused by corrosion and external interference. Reliability Engineering & System Safety. 2020;203.

[14] Yu J, Chen H, Yu Y, Yang Z. A weakest t-norm based fuzzy fault tree approach for leakage risk assessment of submarine pipeline. Journal of Loss Prevention in the Process Industries. 2019;62.

[15] Jabbari M, Gholamnia R, Esmaeili R, Kouhpaee H, Pourtaghi G. Risk assessment of fire, explosion and release of toxic gas of Siri-Assalouyeh sour gas pipeline using fuzzy analytical hierarchy process. Heliyon. 2021;7.

[16] Alves SDT, Lima GBA. Establishing an onshore pipeline incident database to support operational risk management in Brazil-Part 2: Bowtie proposition and statistics of failure. Process Safety and Environmental Protection. 2021;155:80-97.

[17] Centre NT. Risk and emergency preparedness analysis, NORSOK STANDARD Z-013 2001.

[18] STANDARD B. Risk management— Principles and guidelines, BS ISO 31000:2009 2010.

[19] Khan F, Yarveisy R, Abbassi R. Risk-based pipeline integrity management: A road map for the resilient pipelines. Journal of Pipeline Science and Engineering. 2021;1:74-87.

[20] Godaliyadde D. Application of Formal Safety Assessment for Ship Hull Vibration

Modelling [Ph.D.]. Ann Arbor: Liverpool John Moores University (United Kingdom); 2008.

[21] Li X, Zhang Y, Abbassi R, Yang M, Zhang R, Chen G. Dynamic probability assessment of urban natural gas pipeline accidents considering integrated external activities. Journal of Loss Prevention in the Process Industries. 2021;69:104388.

[22] Wang WH, Mou D, Li F, Dong CAF, Khan F. Dynamic failure probability analysis of urban gas pipeline network. Journal of Loss Prevention in the Process Industries. 2021;72.

[23] Lu H, Guo L, Azimi M, Huang K. Oil and Gas 4.0 era: A systematic review and outlook. Computers in Industry. 2019;111:68-90.

[24] Hanga KM, Kovalchuk Y. Machine learning and multi-agent systems in oil and gas industry applications: A survey. Computer Science Review. 2019;34:100191.

[25] Mohammadpoor M, Torabi F. Big Data analytics in oil and gas industry: An emerging trend. Petroleum. 2020;6:321-8.

[26] Mahmoodian M. Chapter 3 - Methods for Structural Reliability Analysis. In: Mahmoodian M, editor. Reliability and Maintainability of In-Service Pipelines: Gulf Professional Publishing; 2018. p. 79-101.

[27] Li X, Wang J, Abbassi R, Chen G. A risk assessment framework considering uncertainty for corrosion-induced natural gas pipeline accidents. Journal of Loss Prevention in the Process Industries. 2022;75:104718.

[28] Joshi AV. Statistical analysis of in-line inspection performance with Gamma distribution. CORROSION 2011: OnePetro; 2011.

[29] Wen H, Khan F, Amin MT, Halim SZ. Myths and misconceptions of data-driven methods: Applications to process safety analysis. Computers & Chemical Engineering. 2022;158:107639.

[30] Baek J-h, Jang Y-c, Kim I-j, Yoo J-s, Kim C-m, Kim Y-p. Structural reliability analysis of in-service API X65 natural gas pipeline using statistical data. International Journal of Pressure Vessels and Piping. 2022;199:104699.

[31] Yu JX, Wu SB, Chen HC, Yu Y, Fan HZ, Liu JH. Risk assessment of submarine

pipelines using modified FMEA approach based on cloud model and extended VIKOR method. Process Safety and Environmental Protection. 2021;155:555-74.

[32] Hemati M, Nenmatpour M, Amini J, Saborifard M. Using failure modes and effects analysis (FMEA) to risk assessment pipelines of petroleum products. Journal of Indian Management & Strategy. 2017;22:4-9.

[33] Zheng DF. Application of HAZOP Study in Key SOP of Oil and Gas Pipelines. International Conference on Solar Energy Materials and Energy Engineering (SEMEE 2013). Hong Kong, PEOPLES R CHINA. 2013. p. 112-7.

[34] Wang W, Ma DH, Su JY, Han Y, Guo XD, Wang ZT. Long Distance Pipeline Reliability Analysis Model Based on Set Pair Analysis and Fault Tree Analysis. International Symposium on Safety Science and Technology. Beijing, PEOPLES R CHINA. 2008. p. 64-7.

[35] Hidayat MEN, Hermansyah H. Risk Analysis of Gas Distribution Pipelines in Apartments using the Event Tree Analysis Method. 8th Annual Basic Science International Conference (BaSIC) - Coverage of Basic Sciences toward the World's Sustainability Challanges. Malang, INDONESIA. 2018.

[36] Hao MJ, You QJ, Yue Z. Risk analysis of urban gas pipeline network based on improved bow-tie model. 2nd International Conference on New Energy and Future Energy System (NEFES). Kunming, PEOPLES R CHINA. 2017.

[37] Muniz MVP, Lima GBA, Caiado RGG, Quelhas OLG. Bow Tie to Improve Risk Management of Natural Gas Pipelines. Process Safety Progress. 2018;37:169-75.

[38] Taleb-Berrouane M, Khan F, Hawboldt K. Corrosion risk assessment using adaptive bow-tie (ABT) analysis. Reliability Engineering & System Safety. 2021;214.

[39] Guo Y, Meng X, Wang D, Meng T, Liu S, He R. Comprehensive risk evaluation of long-distance oil and gas transportation pipelines using a fuzzy Petri net model. Journal of Natural Gas Science and Engineering. 2016;33:18-29.

[40] He S, Xu H, Zhang J, Xue P. Risk assessment of oil and gas pipelines hot work based on AHP-FCE. Petroleum. 2022.

[41] Cui Y, Quddus N, Mashuga CV. Bayesian network and game theory risk

assessment model for third-party damage to oil and gas pipelines. Process Safety and Environmental Protection. 2020;134:178-88.

[42] Flage R, Aven T, Zio E, Baraldi P. Concerns, Challenges, and Directions of Development for the Issue of Representing Uncertainty in Risk Assessment. Risk Analysis. 2014;34:1196-207.

[43] Aven T. On the Need for Restricting the Probabilistic Analysis in Risk Assessments to Variability. Risk Analysis. 2010;30:354-60.

[44] Aven T, Zio E. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliability Engineering & System Safety. 2011;96:64-74.

[45] Yazdi M, Kabir S, Walker M. Uncertainty handling in fault tree based risk assessment: State of the art and future perspectives. Process Safety and Environmental Protection. 2019;131:89-104.

[46] Badida P, Balasubramaniam Y, Jayaprakash J. Risk evaluation of oil and natural gas pipelines due to natural hazards using fuzzy fault tree analysis. Journal of Natural Gas Science and Engineering. 2019;66:284-92.

[47] Dawood T, Elwakil E, Novoa HM, Delgado JFG. Soft computing for modeling pipeline risk index under uncertainty. Engineering Failure Analysis. 2020;117:104949.

[48] Dong Y, Li C-C, Herrera F. Connecting the linguistic hierarchy and the numerical scale for the 2-tuple linguistic model and its use to deal with hesitant unbalanced linguistic information. Information Sciences. 2016;367-368:259-78.

[49] Zarei E, Azadeh A, Khakzad N, Aliabadi MM, Mohammadfam I. Dynamic safety assessment of natural gas stations using Bayesian network. Journal of Hazardous Materials. 2017;321:830-40.

[50] Sarkar S, Quddus N, Mannan MS, El-Halwagi MM. Integrating flare gas with cogeneration systems: Operational risk assessment. Journal of Loss Prevention in the Process Industries. 2021;72:104571.

[51] Idachaba F, Tomomewo O. Surface pipeline leak detection using realtime sensor data analysis. Journal of Pipeline Science and Engineering. 2023;3:100108.

[52] Zuo Z, Ma L, Liang S, Liang J, Zhang H, Liu T. A semi-supervised leakage detection method driven by multivariate time series for natural gas gathering pipeline. Process Safety and Environmental Protection. 2022;164:468-78.
[53] Bhaskaran PE, Chennippan M, Subramaniam T. Future prediction & estimation of faults occurrences in oil pipelines by using data clustering with time series forecasting. Journal of Loss Prevention in the Process Industries. 2020;66:104203.
[54] Li X, Zhao H, Zhang R. Data-driven dynamic failure assessment of subsea gas pipeline using process monitoring data. Process Safety and Environmental Protection. 2022;166:1-10.
[55] Priyanka EB, Thangavel S, Gao X-Z, Sivakumar NS. Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques. Journal of Industrial Information Integration. 2022;26:100272.
[56] Liu C, Wang Y, Li X, Li Y, Khan F, Cai B. Quantitative assessment of leakage orifices within gas pipelines using a Bayesian network. Reliability Engineering & System Safety. 2021;209:107438.
[57] Aljameel SS, Alomari DM, Alismail S, Khawaher F, Alkhudhair AA, Aljubran F, et al. An Anomaly Detection Model for Oil and Gas Pipelines Using Machine Learning. Computation. 2022;10:138.
[58] Lu J, Fu Y, Yue J, Zhu L, Wang D, Hu Z. Natural gas pipeline leak diagnosis based on improved variational modal decomposition and locally linear embedding feature extraction method. Process Safety and Environmental Protection. 2022;164:857-67.
[59] Yang D, Hou N, Lu J, Ji D. Novel leakage detection by ensemble 1DCNN-VAPSO-SVM in oil and gas pipeline systems. Applied Soft Computing. 2022;115:108212.
[60] Li Q, Shi Y, Lin R, Qiao W, Ba W. A novel oil pipeline leakage detection method based on the sparrow search algorithm and CNN. Measurement. 2022;204:112122.
[61] Wang F, Liu Z, Zhou X, Li S, Yuan X, Zhang Y, et al. (INVITED)Oil and Gas Pipeline Leakage Recognition Based on Distributed Vibration and Temperature Information Fusion. Results in Optics. 2021;5:100131.
[62] Guerriero M, Wheeler F, Koste G, Dekate S, Choudhury N. Bayesian data fusion

for pipeline leak detection. 2016 19th International Conference on Information Fusion (FUSION)2016. p. 278-85.

[63] Tennyson RC, Morison WD, Colpitts B, Brown A. Application of Brillouin Fiber Optic Sensors to Monitor Pipeline Integrity. 2004 International Pipeline Conference2004. p. 1259-64.

[64] Fagbami D, Echem C, Okoli A, Mondanos M, Bain A, Carbonneau P, et al. A Practical Application of Pipeline Surveillance and Intrusion Monitoring System in the Niger Delta: The Umugini Case Study. SPE Nigeria Annual International Conference and Exhibition2017.

[65] Wang G, Li X, Li Y. The application of wavelet transform in the fiber pipeline safety pre-warning system. The 27th Chinese Control and Decision Conference (2015 CCDC)2015. p. 3605-8.

[66] Tan D, Tian X, Sun W, Zhou Y, Liu L, Ma Y, et al. An oil and gas pipeline pre-warning system based on Φ-OTDR: SPIE; 2014.

[67] Yan Q, Yang Y, Zhang Z, Jiang S, Yimin X, Zhou C. Distributed vibration sensing system for oil and gas pipelines based on COTDR and BP neural network. Optical Fiber Sensors: Optica Publishing Group; 2018. p. WF42.

[68] Zhao F, Zhang S, Zhao H, Yu L, Feng Q, He J. An intelligent optical fiber-based prewarning system for oil and gas pipelines. Optical Fiber Technology. 2022;71:102953.

[69] Ling J, Zhang H, Dong S, Luo J. Intelligent Prevention Method for Third-Party Damage of Long-Distance Pipeline Based on Mobile Devices Location Information. 2020 13th International Pipeline Conference2020.

[70] Lukowski TI, Power D, Yue B, Randall CJ, Youden J, Howell C. Pipeline encroachment monitoring using polarimetric SAR imagery. IGARSS 2004 2004 IEEE International Geoscience and Remote Sensing Symposium: IEEE; 2004.

[71] Keqin Ding GC, Fangxiong Tang, Li Chen, Yaying He, Liqi Yi. A Study of On-line Monitoring and Evaluation Approach for Oil and Gas Pipelines Subsidence Deformation. 8th European Workshop On Structural Health Monitoring. Spain, Bilbao. e-Journal of Nondestructive Testing; 2016.

[72] Paeper S, Brown B, Beuker T. Inline Inspection of Dents and Corrosion Using “High Quality” Multi-Purpose Smart-Pig Inspection Data. 2006 International Pipeline Conference2006. p. 243-8.

[73] Wenhua H, Peiwan Q. A modified wavelet transform domain adaptive FIR filtering algorithm for removing the SPN contained in the MFL data. 2005 IEEE International Conference on Industrial Technology2005. p. 152-7.

[74] Zhang Y, Ye Z, Xu X. An adaptive method for channel equalization in MFL inspection. NDT & E International. 2007;40:127-39.

[75] Kathirmani S, Tangirala AK, Saha S, Mukhopadhyay S. Online data compression of MFL signals for pipeline inspection. NDT & E International. 2012;50:1-9.

[76] Lang X. Recognition Method of Pipeline Weld Defects Based on Auxiliary Classifier Generative Adversarial Networks. IEEE Instrumentation & Measurement Magazine. 2022;25:69-77.

[77] Witek M. Validation of in-line inspection data quality and impact on steel pipeline diagnostic intervals. Journal of Natural Gas Science and Engineering. 2018;56:121-33.

[78] Carvalho AA, Rebello JMA, Sagrilo LVS, Camerini CS, Miranda IVJ. MFL signals and artificial neural networks applied to detection and classification of pipe weld defects. NDT & E International. 2006;39:661-7.

[79] Feng J, Li F, Lu S, Liu J, Ma D. Injurious or Noninjurious Defect Identification From MFL Images in Pipeline Inspection Using Convolutional Neural Network. IEEE Transactions on Instrumentation and Measurement. 2017;66:1883-92.

[80] Lu S, Feng J, Zhang H, Liu J, Wu Z. An Estimation Method of Defect Size From MFL Image Using Visual Transformation Convolutional Neural Network. IEEE Transactions on Industrial Informatics. 2019;15:213-24.

[81] Joshi A, Udpa L, Udpa S, Tamburrino A. Adaptive Wavelets for Characterizing Magnetic Flux Leakage Signals From Pipeline Inspection. IEEE Transactions on Magnetics. 2006;42:3168-70.

[82] Chen J, Huang S, Zhao W. Three-dimensional defect inversion from magnetic flux leakage signals using iterative neural network. IET Science, Measurement &

Technology. 2015;9:418-26.

[83] Zhongli M, Hongda L. Pipeline defect detection and sizing based on MFL data using immune RBF neural networks. 2007 IEEE Congress on Evolutionary Computation2007. p. 3399-403.

[84] Kandroodi MR, Shirani F, Araabi BN, Ahmadabadi MN, Bassiri MM. Defect detection and width estimation in natural gas pipelines using MFL signals. 2013 9th Asian Control Conference (ASCC)2013. p. 1-6.

[85] Mohamed A, Hamdi MS, Tahar S. A Machine Learning Approach for Big Data in Oil and Gas Pipelines. 2015 3rd International Conference on Future Internet of Things and Cloud2015. p. 585-90.

[86] Layouni M, Hamdi MS, Tahar S. Detection and sizing of metal-loss defects in oil and gas pipelines using pattern-adapted wavelets and machine learning. Applied Soft Computing. 2017;52:247-61.

[87] Wang L, Zhang H, Liu J, Qu F, Zuo F. Defect Size Quantification for Pipeline Magnetic Flux Leakage Detection System via Multilevel Knowledge-Guided Neural Network. IEEE Transactions on Industrial Electronics. 2023;70:9550-60.

[88] Yuksel V, Tetik YE, Basturk MO, Recepoglu O, Gokce K, Cimen MA. A Novel Cascaded Deep Learning Model for the Detection and Quantification of Defects in Pipelines via Magnetic Flux Leakage Signals. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-9.

[89] Piao G, Guo J, Hu T, Deng Y, Leung H. A novel pulsed eddy current method for high-speed pipeline inline inspection. Sensors and Actuators A: Physical. 2019;295:244-58.

[90] Xie M, Tian Z. A review on pipeline integrity management utilizing in-line inspection data. Engineering Failure Analysis. 2018;92:222-39.

[91] Klann M, Beuker T. Pipeline Inspection With the High Resolution EMAT ILI-Tool: Report on Full-Scale Testing and Field Trials. 2006 International Pipeline Conference2006. p. 235-41.

[92] Pyle RJ, Hughes RR, Wilcox PD. Interpretable and Explainable Machine Learning

for Ultrasonic Defect Sizing. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. 2023;70:277-90.

[93] Xiong J, Liang W, Liang X, Yao J. Intelligent quantification of natural gas pipeline defects using improved sparrow search algorithm and deep extreme learning machine. Chemical Engineering Research and Design. 2022;183:567-79.

[94] Yan Y, Liu D, Gao B, Tian GY, Cai ZC. A Deep Learning-Based Ultrasonic Pattern Recognition Method for Inspecting Girth Weld Cracking of Gas Pipeline. IEEE Sensors Journal. 2020;20:7997-8006.

[95] Zhang H, Dong S, Ling J, Zhang L, Cheang B. A modified method for the safety factor parameter: The use of big data to improve petroleum pipeline reliability assessment. Reliability Engineering & System Safety. 2020;198:106892.

[96] Xiang W, Zhou W. Integrated pipeline corrosion growth modeling and reliability analysis using dynamic Bayesian network and parameter learning technique. Structure and Infrastructure Engineering. 2020;16:1161-76.

[97] Zhang H, Tian Z. Reliability assessment of corroded pipeline considering multiple defects interaction based on an artificial neural network method. 2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM)2020. p. 1-6.

[98] Anghel CI. Risk assessment for pipelines with active defects based on artificial intelligence methods. International Journal of Pressure Vessels and Piping. 2009;86:403-11.

[99] Adumene S, Khan F, Adedigba S, Zendehboudi S. Offshore system safety and reliability considering microbial influenced multiple failure modes and their interdependencies. Reliability Engineering & System Safety. 2021;215:107862.

[100] Shabarchin O, Tesfamariam S. Internal corrosion hazard assessment of oil & gas pipelines using Bayesian belief network model. Journal of Loss Prevention in the Process Industries. 2016;40:479-95.

[101] Woldesellasse H, Tesfamariam S. Failure assessment of oil and gas transmission pipelines using an integrated Bayesian belief network and GIS model. International

Journal of Pressure Vessels and Piping. 2023;205:104984.
[102] Mandal K, Dufour D, Atherton DL. Use of magnetic Barkhausen noise and magnetic flux leakage signals for analysis of defects in pipeline steel. IEEE Transactions on Magnetics. 1999;35:2007-17.
[103] Guo X, Zhang L, Liang W, Haugen S. Risk identification of third-party damage on oil and gas pipelines through the Bayesian network. Journal of Loss Prevention in the Process Industries. 2018;54:163-78.
[104] Wen H, Liu L, Zhang J, Hu J, Huang X. A hybrid machine learning model for landslide-oriented risk assessment of long-distance pipelines. Journal of Environmental Management. 2023;342:118177.
[105] Shan K, Shuai J, Xu K, Zheng W. Failure probability assessment of gas transmission pipelines based on historical failure-related data and modification factors. Journal of Natural Gas Science and Engineering. 2018;52:356-66.
[106] Mazumder RK, Salman AM, Li Y. Failure risk analysis of pipelines using data-driven machine learning algorithms. Structural Safety. 2021;89:102047.
[107] Yang Y, Li S, Zhang P. Data-driven accident consequence assessment on urban gas pipeline network based on machine learning. Reliability Engineering & System Safety. 2022;219:108216.
[108] Kumari P, Halim SZ, Kwon JS-I, Quddus N. An integrated risk prediction model for corrosion-induced pipeline incidents using artificial neural network and Bayesian analysis. Process Safety and Environmental Protection. 2022;167:34-44.
[109] Alves DTS, Lima GBA. Establishing an onshore pipeline incident database to support operational risk management in Brazil - Part 2: Bowtie proposition and statistics of failure. Process Safety and Environmental Protection. 2021;155:80-97.
[110] Wang Z, Li S. Data-driven risk assessment on urban pipeline network based on a cluster model. Reliability Engineering & System Safety. 2020;196:106781.
[111] Ma H, Wang H, Geng M, Ai Y, Zhang W, Zheng W. A new hybrid approach model for predicting burst pressure of corroded pipelines of gas and oil. Engineering Failure Analysis. 2023;149:107248.

[112] Li X, Wang J, Chen G. A machine learning methodology for probabilistic risk assessment of process operations: A case of subsea gas pipeline leak accidents. Process Safety and Environmental Protection. 2022;165:959-68.
[113] Hassan S, Wang J, Kontovas C, Bashir M. An assessment of causes and failure likelihood of cross-country pipelines under uncertainty using bayesian networks. Reliability Engineering & System Safety. 2022;218:108171.
[114] Feng X, Jiang J-c, Wang W-f. Gas pipeline failure evaluation method based on a Noisy-OR gate bayesian network. Journal of Loss Prevention in the Process Industries. 2020;66:104175.
[115] Raeihagh H, Behbahaninia A, Macki Aleagha M. Risk assessment of sour gas inter-phase onshore pipeline using ANN and fuzzy inference system – Case study: The south pars gas field. Journal of Loss Prevention in the Process Industries. 2020;68:104238.
[116] Su Y, Li J, Yu B, Zhao Y, Yao J. Fast and accurate prediction of failure pressure of oil and gas defective pipelines using the deep learning model. Reliability Engineering & System Safety. 2021;216:108016.
[117] Alves DTS, Lima GBA. Establishing an onshore pipeline incident database to support operational risk management in Brazil - part 1: Defining architecture. Process Safety and Environmental Protection. 2021;154:480-504.
[118] Pasman HJ, Rogers WJ. How to treat expert judgment? With certainty it contains uncertainty! Journal of Loss Prevention in the Process Industries. 2020;66:104200.
[119] Shahriar A, Sadiq R, Tesfamariam S. Risk analysis for oil & gas pipelines: A sustainability assessment approach using fuzzy based bow-tie analysis. Journal of Loss Prevention in the Process Industries. 2012;25:505-23.
[120] Yuhua D, Datao Y. Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis. Journal of Loss Prevention in the Process Industries. 2005;18:83-8.
[121] Jianxing Y, Haicheng C, Yang Y, Zhenglong Y. A weakest t-norm based fuzzy fault tree approach for leakage risk assessment of submarine pipeline. Journal of Loss

Prevention in the Process Industries. 2019;62:103968.
[122] Yu J, Zeng Q, Yu Y, Wu S, Ding H, Gao H, et al. An intuitionistic fuzzy probabilistic Petri net method for risk assessment on submarine pipeline leakage failure. Ocean Engineering. 2022;266:112788.
[123] Lu L, Liang W, Zhang L, Zhang H, Lu Z, Shan J. A comprehensive risk evaluation method for natural gas pipelines by combining a risk matrix with a bow-tie model. Journal of Natural Gas Science and Engineering. 2015;25:124-33.
[124] Hong B, Shao B, Guo J, Fu J, Li C, Zhu B. Dynamic Bayesian network risk probability evolution for third-party damage of natural gas pipelines. Applied Energy. 2023;333:120620.
[125] Wang W, He X, Li Y, Shuai J. Risk analysis on corrosion of submarine oil and gas pipelines based on hybrid Bayesian network. Ocean Engineering. 2022;260:111957.
[126] Hu J, Chen C, Liu Z. Early warning method for overseas natural gas pipeline accidents based on FDOOBN under severe environmental conditions. Process Safety and Environmental Protection. 2022;157:175-92.
[127] Chen X-l, Lin W-d, Liu C-x, Yang F-q, Guo Y, Li X, et al. An integrated EDIB model for probabilistic risk analysis of natural gas pipeline leakage accidents. Journal of Loss Prevention in the Process Industries. 2023;83:105027.
[128] Yuan C, Hu Y, Zhang Y, Zuo T, Wang J, Fan S. Evaluation on consequences prediction of fire accident in emergency processes for oil-gas storage and transportation by scenario deduction. Journal of Loss Prevention in the Process Industries. 2021;72:104570.
[129] Yin H, Liu C, Wu W, Song K, Dan Y, Cheng G. An integrated framework for criticality evaluation of oil & gas pipelines based on fuzzy logic inference and machine learning. Journal of Natural Gas Science and Engineering. 2021;96:104264.
[130] Chen Y, Xie S, Tian Z. Risk assessment of buried gas pipelines based on improved cloud-variable weight theory. Reliability Engineering & System Safety. 2022;221:108374.
[131] Li X, Han Z, Yazdi M, Chen G. A CRITIC-VIKOR based robust approach to

support risk management of subsea pipelines. Applied Ocean Research. 2022;124:103187.

[132] Jianxing Y, Shibo W, Haicheng C, Yang Y, Haizhao F, Jiahao L. Risk assessment of submarine pipelines using modified FMEA approach based on cloud model and extended VIKOR method. Process Safety and Environmental Protection. 2021;155:555-74.

[133] Hassan S, Wang J, Kontovas C, Bashir M. Modified FMEA hazard identification for cross-country petroleum pipeline using Fuzzy Rule Base and approximate reasoning. Journal of Loss Prevention in the Process Industries. 2022;74:104616.

[134] Guo Y, Meng X, Meng T, Wang D, Liu S. A novel method of risk assessment based on cloud inference for natural gas pipelines. Journal of Natural Gas Science and Engineering. 2016;30:421-9.

[135] Li X, Han Z, Zhang R, Abbassi R, Chang D. An integrated methodology to manage risk factors of aging urban oil and gas pipelines. Journal of Loss Prevention in the Process Industries. 2020;66:104154.

[136] Chen Y, Hou F, Dong S, Guo L, Xia T, He G. Reliability evaluation of corroded pipeline under combined loadings based on back propagation neural network method. Ocean Engineering. 2022;262:111910.

[137] Zhang T, Shuai J, Shuai Y, Hua L, Xu K, Xie D, et al. Efficient prediction method of triple failure pressure for corroded pipelines under complex loads based on a backpropagation neural network. Reliability Engineering & System Safety. 2023;231:108990.

[138] Xu D, Chen L, Yu C, Zhang S, Zhao X, Lai X. Failure analysis and control of natural gas pipelines under excavation impact based on machine learning scheme. International Journal of Pressure Vessels and Piping. 2023;201:104870.

[139] Peng X-y, Yao D-c, Liang G-c, Yu J-s, He S. Overall reliability analysis on oil/gas pipeline under typical third-party actions based on fragility theory. Journal of Natural Gas Science and Engineering. 2016;34:993-1003.

[140] Jiang F, Dong S. Collision failure risk analysis of falling object on subsea

pipelines based on machine learning scheme. Engineering Failure Analysis. 2020;114:104601.

[141] Ossai CI, Boswell B, Davies IJ. Estimation of Internal Pit Depth Growth and Reliability of Aged Oil and Gas Pipelines— A Monte Carlo Simulation Approach. Corrosion. 2015;71:977-91.

[142] Timashev SA, Malyukova MG, Poluian LV, Bushinskaya AV. Markov Description of Corrosion Defects Growth and Its Application to Reliability Based Inspection and Maintenance of Pipelines. 2008 7th International Pipeline Conference2008. p. 525-33.

[143] Park K, Lee G, Nam C, Lee W. Quantitative and Probabilistic approach for Underground Pipeline management Optimization. In: Eden MR, Ierapetritou MG, Towler GP, editors. Computer Aided Chemical Engineering: Elsevier; 2018. p. 1549-54.

[144] Bandstra D, Fraser AM. Subset Simulation for Structural Reliability Analysis of Pipeline Corrosion Defects. 2020 13th International Pipeline Conference2020.

[145] Chakraborty S, Tesfamariam S. Subset simulation based approach for space-time-dependent system reliability analysis of corroding pipelines. Structural Safety. 2021;90:102073.

[146] Seghier MEAB, Spyridis P, Jafari-Asl J, Ohadi S, Li X. Comparative Study on the Efficiency of Simulation and Meta-Model-Based Monte Carlo Techniques for Accurate Reliability Analysis of Corroded Pipelines. Sustainability. 2022;14:5830.

[147] Pasman H, Rogers W. How trustworthy are risk assessment results, and what can be done about the uncertainties they are plagued with? Journal of Loss Prevention in the Process Industries. 2018;55:162-77.

[148] Fenyvesi LL, Colquhoun I, Kania R, Gu B. A Risk-Based Approach to Maintenance Planning Utilizing In-Line Inspection Data. 2004 International Pipeline Conference 2004. p. 2609-19.

[149] Salama MM, Nestleroth BJ, Maes MA, Dash C. Characterization of the Uncertainties in the Inspection Results of Ultrasonic Intelligent Pigs. ASME 2013 32nd

International Conference on Ocean, Offshore and Arctic Engineering, 2013.
[150] Singh M, Markeset T. Hybrid models for handling variability and uncertainty in probabilistic and possibilistic failure analysis of corroded pipes. Engineering Failure Analysis. 2014;42:197-209.
[151] Wu W, Li Y, Cheng G, Zhang H, Kang J. Dynamic safety assessment of oil and gas pipeline containing internal corrosion defect using probability theory and possibility theory. Engineering Failure Analysis. 2019;98:156-66.
[152] Dann MR, Maes MA. Stochastic corrosion growth modeling for pipelines using mass inspection data. Reliability Engineering & System Safety. 2018;180:245-54.
[153] Niu Z, Zhu H, Huang X, Che A, Fu S, Meng S, et al. Uncertainty quantification method for elastic wave tomography of concrete structure using interval analysis. Measurement. 2022;205:112160.
[154] Yu X, Liang W, Zhang L, Reniers G, Lu L. Risk assessment of the maintenance process for onshore oil and gas transmission pipelines under uncertainty. Reliability Engineering & System Safety. 2018;177:50-67.
[155] Zio E, Pedroni N. Literature review of methods for representing uncertainty. 2013.
[156] Singh K, Kaushik M, Kumar M. Integrating α-cut interval based fuzzy fault tree analysis with Bayesian network for criticality analysis of submarine pipeline leakage: A novel approach. Process Safety and Environmental Protection. 2022;166:189-201.
[157] Xie S, Dong S, Chen Y, Peng Y, Li X. A novel risk evaluation method for fire and explosion accidents in oil depots using bow-tie analysis and risk matrix analysis method based on cloud model theory. Reliability Engineering & System Safety. 2021;215:107791.
[158] Wang Y, Hou X, Zhang P, Qin G. Reliability assessment of multi-state reconfiguration pipeline system with failure interaction based on Cloud inference. Process Safety and Environmental Protection. 2020;137:116-27.
[159] Liang X, Liang W, Zhang L, Guo X. Risk assessment for long-distance gas pipelines in coal mine gobs based on structure entropy weight method and multi-step backward cloud transformation algorithm based on sampling with replacement. Journal

of Cleaner Production. 2019;227:218-28.

[160] Bisaggio HdC, Netto TA. Predictive analyses of the integrity of corroded pipelines based on concepts of structural reliability and Bayesian inference. Marine Structures. 2015;41:180-99.

[161] Zhang S, Zhou W. Bayesian dynamic linear model for growth of corrosion defects on energy pipelines. Reliability Engineering & System Safety. 2014;128:24-31.

[162] Li X, Liu Y, Abbassi R, Khan F, Zhang R. A Copula-Bayesian approach for risk assessment of decommissioning operation of aging subsea pipelines. Process Safety and Environmental Protection. 2022;167:412-22.

[163] Qin H, Zhou W, Zhang S. Bayesian inferences of generation and growth of corrosion defects on energy pipelines based on imperfect inspection data. Reliability Engineering & System Safety. 2015;144:334-42.

[164] Pesinis K, Tee KF. Bayesian analysis of small probability incidents for corroding energy pipelines. Engineering Structures. 2018;165:264-77.

[165] Heidary R, Groth KM. A hybrid population-based degradation model for pipeline pitting corrosion. Reliability Engineering & System Safety. 2021;214:107740.

[166] Liu A, Chen K, Huang X, Li D, Zhang X. Dynamic risk assessment model of buried gas pipelines based on system dynamics. Reliability Engineering & System Safety. 2021;208:107326.

[167] Taleb-Berrouane M, Khan F, Hawboldt K. Corrosion risk assessment using adaptive bow-tie (ABT) analysis. Reliability Engineering & System Safety. 2021;214:107731.

[168] Li X, Chen G, Zhu H. Quantitative risk analysis on leakage failure of submarine oil and gas pipelines using Bayesian network. Process Safety and Environmental Protection. 2016;103:163-73.

[169] Dundulis G, Žutautaitė I, Janulionis R, Uškuras E, Rimkevičius S, Eid M. Integrated failure probability estimation based on structural integrity analysis and failure data: Natural gas pipeline case. Reliability Engineering & System Safety. 2016;156:195-202.

[170] Park KS, Lee JH, Jo YD. An Approach to Risk Management of City Gas Pipeline. Process Safety and Environmental Protection. 2004;82:446-52.

[171] Vairo T, Pontiggia M, Fabiano B. Critical aspects of natural gas pipelines risk assessments. A case-study application on buried layout. Process Safety and Environmental Protection. 2021;149:258-68.

[172] Li X, Yang M, Chen G. An integrated framework for subsea pipelines safety analysis considering causation dependencies. Ocean Engineering. 2019;183:175-86.

[173] Li F, Wang W, Xu J, Yi J, Wang Q. Comparative study on vulnerability assessment for urban buried gas pipeline network based on SVM and ANN methods. Process Safety and Environmental Protection. 2019;122:23-32.

[174] Qiu Z, Liang W, Zhang L. Tracing and prediction analysis of an urban pipeline leakage accident based on the catastrophe DBN model. Journal of Natural Gas Science and Engineering. 2018;57:339-48.

[175] Luo T, Wu C, Duan L. Fishbone diagram and risk matrix analysis method and its application in safety assessment of natural gas spherical tank. Journal of Cleaner Production. 2018;174:296-304.

[176] Karimi H, Sadeghi-Dastaki M, Javan M. A fully fuzzy best–worst multi attribute decision making method with triangular fuzzy number: A case study of maintenance assessment in the hospitals. Applied Soft Computing. 2020;86:105882.

[177] Xie S, Chen Y, Dong S, Zhang G. Risk assessment of an oil depot using the improved multi-sensor fusion approach based on the cloud model and the belief Jensen-Shannon divergence. Journal of Loss Prevention in the Process Industries. 2020;67:104214.

[178] Peng T, Deng H, Lin Y, Jin Z. Assessment on water resources carrying capacity in karst areas by using an innovative DPESBRM concept model and cloud model. Science of The Total Environment. 2021;767:144353.

[179] Qin G, Zhang M, Yan Q, Xu C, Kammen DM. Comprehensive evaluation of regional energy internet using a fuzzy analytic hierarchy process based on cloud model: A case in China. Energy. 2021;228:120569.

[180] Ruan D, Bian J, Wang Q, Wu J, Yu Y, Gu Z. Application of Modified Cloud Model-Level Eigenvalue Method in water quality evaluation. Journal of Hydrology. 2021;603:126980.

[181] Yao J, Wang G, Wang L, Zhang X, Wang P. Assessing the spatiotemporal variability of lake water quality using A novel multidimensional shape – Position similarity cloud model. Journal of Hydrology. 2021;599:126379.

[182] Li D, Liu C, Gan W. A new cognitive model: Cloud model. International Journal of Intelligent Systems. 2009;24:357-75.

[183] Cao Y, Bian Y. Improving the ecological environmental performance to achieve carbon neutrality: The application of DPSIR-Improved matter-element extension cloud model. Journal of Environmental Management. 2021;293:112887.

[184] Yan Q, Zhang M, Li W, Qin G. Risk Assessment of New Energy Vehicle Supply Chain Based on Variable Weight Theory and Cloud Model: A Case Study in China. Sustainability. 2020;12:3150.

[185] Liu S, Li W, Wang Q. Zoning method for environmental engineering geological patterns in underground coal mining areas. Science of The Total Environment. 2018;634:1064-76.

[186] Lin C, Zhang M, Zhou Z, Li L, Shi S, Chen Y, et al. A new quantitative method for risk assessment of water inrush in karst tunnels based on variable weight function and improved cloud model. Tunnelling and Underground Space Technology. 2020;95:103136.

[187] Zheng G, Wang Y, Li C, Wang X. Real-time quantification of human physiological state in high temperature environments based on variable weight theory. Journal of Thermal Biology. 2020;89:102531.

[188] Mete S, Serin F, Oz NE, Gul M. A decision-support system based on Pythagorean fuzzy VIKOR for occupational risk assessment of a natural gas pipeline construction. Journal of Natural Gas Science and Engineering. 2019;71:102979.

[189] Boulomytis VTG, Zuffo AC, Imteaz MA. Detection of flood influence criteria in ungauged basins on a combined Delphi-AHP approach. Operations Research

Perspectives. 2019;6:100116.
[190] Yang X, Yan L, Zeng L. How to handle uncertainties in AHP: The Cloud Delphi hierarchical analysis. Information Sciences. 2013;222:384-404.
[191] Abrahamsen EB, Abrahamsen HB, Milazzo MF, Selvik JT. Using the ALARP principle for safety management in the energy production sector of chemical industry. Reliability Engineering & System Safety. 2018;169:160-5.
[192] Mottahedi A, Sereshki F, Ataei M, Qarahasanlou AN, Barabadi A. Resilience estimation of critical infrastructure systems: Application of expert judgment. Reliability Engineering & System Safety. 2021;215:107849.
[193] Yazdi M, Kabir S. A fuzzy Bayesian network approach for risk analysis in process industries. Process Safety and Environmental Protection. 2017;111:507-19.
[194] Yazdi M, Hafezi P, Abbassi R. A methodology for enhancing the reliability of expert system applications in probabilistic risk assessment. Journal of Loss Prevention in the Process Industries. 2019;58:51-9.
[195] Moktadir MA, Dwivedi A, Khan NS, Paul SK, Khan SA, Ahmed S, et al. Analysis of risk factors in sustainable supply chain management in an emerging economy of leather industry. Journal of Cleaner Production. 2021;283:124641.
[196] Lv X, Ding H. A Method of Optimal Malfunction Management in Urban Natural Gas Transmission and Distribution Systems. 4th International Workshop on Renewable Energy and Development (IWRED) 2020.
[197] Nikbakht M, Sayyah A, Zulkifli N. Hazard identification and accident analysis on city gate station in natural gas industry. 2010 International Conference on Environmental Engineering and Applications2010. p. 13-7.
[198] Guo Y, Yang X, Liu B, Liu S. Fracture failure analysis of valve bolt in the gas transmission station. IEEE International Conference on Mechatronics and Automation. Harbin, China. 2016. p. 1310-5.
[199] Wang D, Liang P, Yu Y, Fu X, Hu L. An integrated methodology for assessing accident probability of natural gas distribution station with data uncertainty. Journal of Loss Prevention in the Process Industries. 2019;62.

[200] Incorporated WE. Pipeline Investigation Report P12H0103. Canada TSBo. Available from: www.bst-tsb.gc.ca. September 2013.(accessed 7 April 2022)

[201] Talk:List of pipeline accidents. Available from: https://en.wikipedia.org/wiki/Talk:List_of_pipeline_accidents: WikiProject Lists.(accessed 7 April 2022)

[202] Cezar-Vaz MR, Rocha LP, Bonow CA, Santos da Silva MR, Vaz JC, Cardoso LS. Risk Perception and Occupational Accidents: A Study of Gas Station Workers in Southern Brazil. International Journal of Environmental Research and Public Health. 2012;9:2362-77.

[203] Dadkani P, Noorzai E, Ghanbari A, Gharib A. Risk analysis of gas leakage in gas pressure reduction station and its consequences: A case study for Zahedan. Heliyon. 2021;7.

[204] Skogdalen JE, Vinnem JE. Combining precursor incidents investigations and QRA in oil and gas industry. Reliability Engineering & System Safety. 2012;101:48-58.

[205] Saleh JH, Saltmarsh EA, Favaro FM, Brevault L. Accident precursors, near misses, and warning signs: Critical review and formal definitions within the framework of Discrete Event Systems. Reliability Engineering & System Safety. 2013;114:148-54.

[206] Liu C, Wang Y, Li X, Li Y, Khan F, Cai B. Quantitative assessment of leakage orifices within gas pipelines using a Bayesian network. Reliability Engineering & System Safety. 2021;209.

[207] Xing L, Sela L. Unsteady pressure patterns discovery from high-frequency sensing in water distribution systems. Water Research. 2019;158:291-300.

[208] BahooToroody A, Abaei MM, BahooToroody F, De Carlo F, Abbassi R, Khalaj S. A condition monitoring based signal filtering approach for dynamic time dependent safety assessment of natural gas distribution process. Process Safety and Environmental Protection. 2019;123:335-43.

[209] Adedigba SA, Khan F, Yang M. Dynamic Failure Analysis of Process Systems Using Principal Component Analysis and Bayesian Network. Industrial & Engineering Chemistry Research. 2017;56:2094-106.

[210] Amin MT, Khan F, Ahmed S, Imtiaz S. A novel data-driven methodology for fault detection and dynamic risk assessment. Canadian Journal of Chemical Engineering. 2020;98:2397-416.

[211] Rebello S, Yu H, Ma L. An integrated approach for real-time hazard mitigation in complex industrial processes. Reliability Engineering & System Safety. 2019;188:297-309.

[212] Leoni L, BahooToroody A, De Carlo F, Paltrinieri N. Developing a risk-based maintenance model for a Natural Gas Regulating and Metering Station using Bayesian Network. Journal of Loss Prevention in the Process Industries. 2019;57:17-24.

[213] Sarkar S, Quddus N, Mannan MS, El-Halwagi MM. Integrating flare gas with cogeneration systems: Operational risk assessment. Journal of Loss Prevention in the Process Industries. 2021;72.

[214] Tan Q, Chen G, Zhang L, Fu J, Li Z. Dynamic accident modeling for high-sulfur natural gas gathering station. Process Safety and Environmental Protection. 2014;92:565-76.

[215] Yun G, Rogers WJ, Mannan MS. Risk assessment of LNG importation terminals using the Bayesian-LOPA methodology. Journal of Loss Prevention in the Process Industries. 2009;22:91-6.

[216] Huang Y, Ma G. A grid-based risk screening method for fire and explosion events of hydrogen refuelling stations. International Journal of Hydrogen Energy. 2018;43:442-54.

[217] Nourian R, Mousavi SM, Raissi S. A fuzzy expert system for mitigation of risks and effective control of gas pressure reduction stations with a real application. Journal of Loss Prevention in the Process Industries. 2019;59:77-90.

[218] Gao P, Li W, Sun Y, Liu S. Risk assessment for gas transmission station based on cloud model based multilevel Bayesian network from the perspective of multi-flow intersecting theory. Process Safety and Environmental Protection. 2022;159:887-98.

[219] Liu S, Li W, Gao P, Sun Y. Modeling and performance analysis of gas leakage emergency disposal process in gas transmission station based on Stochastic Petri nets.

Reliability Engineering & System Safety. 2022;226:108708.
[220] Xu S, Kim E, Haugen S, Zhang M. A Bayesian network risk model for predicting ship besetting in ice during convoy operations along the Northern Sea Route. Reliability Engineering & System Safety. 2022;223:8475.
[221] Chang Y, Wu X, Zhang C, Chen G, Liu X, Li J, et al. Dynamic Bayesian networks based approach for risk analysis of subsea wellhead fatigue failure during service life. Reliability Engineering & System Safety. 2019;188:454-62.
[222] Chang Y, Zhang C, Shi J, Li J, Zhang S, Chen G. Dynamic Bayesian network based approach for risk analysis of hydrogen generation unit leakage. International Journal of Hydrogen Energy. 2019;44:26665-78.
[223] Abimbola M, Khan F. Resilience modeling of engineering systems using dynamic object-oriented Bayesian network approach. Computers & Industrial Engineering. 2019;130:108-18.
[224] Li T, Zhou Y, Zhao Y, Zhang C, Zhang X. A hierarchical object oriented Bayesian network-based fault diagnosis method for building energy systems. Applied Energy. 2022;306.
[225] Liu Q, Tchangani A, Peres F, Gonzalez-Prida V. Object-oriented Bayesian network for complex system risk assessment. Proceedings of the Institution of Mechanical Engineers Part O-Journal of Risk and Reliability. 2018;232:340-51.
[226] Cai B, Liu H, Xie M. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. Mechanical Systems and Signal Processing. 2016;80:31-44.
[227] Yuan X, Cai B, Ma Y, Zhang J, Mulenga K, Liu Y, et al. Reliability Evaluation Methodology of Complex Systems Based on Dynamic Object-Oriented Bayesian Networks. Ieee Access. 2018;6:11289-300.
[228] Abbasnezhad K, Ansari R, Mahdikhani M. Schedule Risk Assessments Using a Precedence Network: An Object-Oriented Bayesian Approach. Iranian Journal of Science and Technology-Transactions of Civil Engineering. 2022;46:1737-53.
[229] Bin Shams MA, Budman HM, Duever TA. Fault detection, identification and

diagnosis using CUSUM based PCA. Chemical Engineering Science. 2011;66:4488-98.

[230] Moradi R, Cofre-Martel S, Droguett EL, Modarres M, Groth KM. Integration of deep learning and Bayesian networks for condition and operation risk monitoring of complex engineering systems. Reliability Engineering & System Safety. 2022;222.

[231] Peeters JFW, Basten RJI, Tinga T. Improving failure analysis efficiency by combining FTA and FMEA in a recursive manner. Reliability Engineering & System Safety. 2018;172:36-44.

[232] Wang Q, Jia G, Jia Y, Song W. A new approach for risk assessment of failure modes considering risk interaction and propagation effects. Reliability Engineering & System Safety. 2021;216.

[233] Chi C-F, Sigmund D, Astardi MO. Classification Scheme for Root Cause and Failure Modes and Effects Analysis (FMEA) of Passenger Vehicle Recalls. Reliability Engineering & System Safety. 2020;200.

[234] van Staalduinen MA, Khan F, Gadag V, Reniers G. Functional quantitative security risk analysis (QSRA) to assist in protecting critical process infrastructure. Reliability Engineering & System Safety. 2017;157:23-34.

[235] Feng X, Jiang J-c, Wang W-f. Gas pipeline failure evaluation method based on a Noisy-OR gate bayesian network. Journal of Loss Prevention in the Process Industries. 2020;66.

[236] Yu J, Wu S, Yu Y, Chen H, Fan H, Liu J, et al. Process system failure evaluation method based on a Noisy-OR gate intuitionistic fuzzy Bayesian network in an uncertain environment. Process Safety and Environmental Protection. 2021;150:281-97.

[237] Crispim J, Fernandes J, Rego N. Customized risk assessment in military shipbuilding. Reliability Engineering & System Safety. 2020;197.

[238] Zhang Z, Guo Y, Wang D, Li G, Peng D. A Noisy-OR gate based fuzzy fault tree approach for micro-leakage evaluation of bolt-gasket-flange connection (BGFC). Journal of Loss Prevention in the Process Industries. 2021;71.

[239] Babaleye AO, Kurt RE, Khan F. Safety analysis of plugging and abandonment of

oil and gas wells in uncertain conditions with limited data. Reliability Engineering & System Safety. 2019;188:133-41.

[240] Huang W, Kou X, Zhang Y, Mi R, Yin D, Xiao W, et al. Operational failure analysis of high-speed electric multiple units: A Bayesian network-K2 algorithm-expectation maximization approach. Reliability Engineering & System Safety. 2021;205.

[241] Li Z, Wang F, Wang C, Hu Q, Yu D. Reliability modeling and evaluation of lifetime delayed degradation process with nondestructive testing. Reliability Engineering & System Safety. 2021;208.

[242] Lee D, Choi D. Analysis of the reliability of a starter-generator using a dynamic Bayesian network. Reliability Engineering & System Safety. 2020;195.

[243] Ren H, Liu M, Li Z, Pedrycz W. A Piecewise Aggregate pattern representation approach for anomaly detection in time series. Knowledge-Based Systems. 2017;135:29-39.

[244] Dao PB. A CUSUM-Based Approach for Condition Monitoring and Fault Diagnosis of Wind Turbines. Energies. 2021;14.

[245] Haller P, Genge B, Duka A-V. On the practical integration of anomaly detection techniques in industrial control applications. International Journal of Critical Infrastructure Protection. 2019;24:48-68.

[246] Raman GMR, Somu N, Mathur AP. A multilayer perceptron model for anomaly detection in water treatment plants. International Journal of Critical Infrastructure Protection. 2020;31.

[247] Kazemi P, Bengoa C, Steyer J-P, Giralt J. Data-driven techniques for fault detection in anaerobic digestion process. Process Safety and Environmental Protection. 2021;146:905-15.

[248] De Oca VM, Jeske DR, Zhang Q, Rendon C, Marvasti M. A cusum change-point detection algorithm for non-stationary sequences with application to data network surveillance. Journal of Systems and Software. 2010;83:1288-97.

[249] Wu Z, Yang M, Jiang W, Khoo MBC. Optimization designs of the combined

Shewhart-CUSUM control charts. Computational Statistics & Data Analysis. 2008;53:496-506.

[250] ISO. Control charts —Part 4: Cumulative sum charts: ISO 7870-4:2021(E); 2021-09.

[251] Liu P, Chen Y, Nie X, Zhu J, Zhang S, Sui K, et al. FluxRank: A Widely-Deployable Framework to Automatically Localizing Root Cause Machines for Software Service Failure Mitigation. 30th IEEE International Symposium on Software Reliability Engineering (ISSRE). Berlin, GERMANY. 2019. p. 35-46.

[252] Liu P, Liu Y, Cai B, Wu X, Wang K, Wei X, et al. A dynamic Bayesian network based methodology for fault diagnosis of subsea Christmas tree. Applied Ocean Research. 2020;94:101990.

[253] He R, Li X, Chen G, Wang Y, Jiang S, Zhi C. A quantitative risk analysis model considering uncertain information. Process Safety and Environmental Protection. 2018;118:361-70.

[254] Wang YF, Qin T, Li B, Sun XF, Li YL. Fire probability prediction of offshore platform based on Dynamic Bayesian Network. Ocean Engineering. 2017;145:112-23.

[255] Jones B, Jenkinson I, Yang Z, Wang J. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliability Engineering & System Safety. 2010;95:267-77.

[256] Ma H, Zhang W, Wang Y, Ai Y, Zheng W. Advances in corrosion growth modeling for oil and gas pipelines: A review. Process Safety and Environmental Protection. 2023;171:71-86.

[257] Dao U, Yarveisy R, Anwar S, Khan F, Zhang Y, Ngo HH. A Bayesian approach to assess under-deposit corrosion in oil and gas pipelines. Process Safety and Environmental Protection. 2023;176:489-505.

[258] Akhlaghi B, Mesghali H, Ehteshami M, Mohammadpour J, Salehi F, Abbassi R. Predictive deep learning for pitting corrosion modeling in buried transmission pipelines. Process Safety and Environmental Protection. 2023;174:320-7.

[259] PHMSA. Gas Transmission HCA Leaks. PHMSA. Available from:

2020.(accessed 18 April 2023)

[260] Valor A, Alfonso L, Caleyo F, Vidal J, Perez-Baruch E, Hallen JM. The negative binomial distribution as a model for external corrosion defect counts in buried pipelines. Corrosion Science. 2015;101:114-31.

[261] Xie M, Tian Z. Risk-based pipeline re-assessment optimization considering corrosion defects. Sustainable Cities and Society. 2018;38:746-57.

[262] Abubakirov R, Yang M, Khakzad N. A risk-based approach to determination of optimal inspection intervals for buried oil pipelines. Process Safety and Environmental Protection. 2020;134:95-107.

[263] Amaya-Gómez R, Sánchez-Silva M, Muñoz F. Integrity assessment of corroded pipelines using dynamic segmentation and clustering. Process Safety and Environmental Protection. 2019;128:284-94.

[264] Wang H, Yajima A, Castaneda H. A stochastic defect growth model for reliability assessment of corroded underground pipelines. Process Safety and Environmental Protection. 2019;123:179-89.

[265] Brockhaus S, Ginten M, Klein S, Teckert M, Stawicki O, Oevermann D, et al. 10 - In-line inspection (ILI) methods for detecting corrosion in underground pipelines. In: Orazem ME, editor. Underground Pipeline Corrosion: Woodhead Publishing; 2014. p. 255-85.

[266] Yang L, Wang Z, Gao S. Pipeline Magnetic Flux Leakage Image Detection Algorithm Based on Multiscale SSD Network. IEEE Transactions on Industrial Informatics. 2020;16:501-9.

[267] Zhang H, Yu X. Research on oil and gas pipeline defect recognition based on IPSO for RBF neural network. Sustainable Computing: Informatics and Systems. 2018;20:203-9.

[268] Ling J, Feng K, Wang T, Liao M, Yang C, Liu Z. Data Modeling Techniques for Pipeline Integrity Assessment: A State-of-the-Art Survey. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-17.

[269] Li X, Jia R, Zhang R, Yang S, Chen G. A KPCA-BRANN based data-driven

approach to model corrosion degradation of subsea oil pipelines. Reliability Engineering & System Safety. 2022;219:108231.

[270] Chen Z, Li X, Wang W, Li Y, Shi L, Li Y. Residual strength prediction of corroded pipelines using multilayer perceptron and modified feedforward neural network. Reliability Engineering & System Safety. 2023;231:108980.

[271] Pourahmadi M, Saybani M. Reliability analysis with corrosion defects in submarine pipeline case study: Oil pipeline in Ab-khark island. Ocean Engineering. 2022;249:110885.

[272] Qin G, Cheng YF, Zhang P. Finite element modeling of corrosion defect growth and failure pressure prediction of pipelines. International Journal of Pressure Vessels and Piping. 2021;194:104509.

[273] Yarveisy R, Khan F, Abbassi R. Data-driven predictive corrosion failure model for maintenance planning of process systems. Computers & Chemical Engineering. 2022;157:107612.

[274] Yu W, Huang W, Wen K, Zhang J, Liu H, Wang K, et al. Subset simulation-based reliability analysis of the corroding natural gas pipeline. Reliability Engineering & System Safety. 2021;213:107661.

[275] Pintelas E, Livieris IE, Kotsiantis S, Pintelas P. A multi-view-CNN framework for deep representation learning in image classification. Computer Vision and Image Understanding. 2023;232:103687.

[276] Liu P, Xu C, Xie J, Fu M, Chen Y, Liu Z, et al. A CNN-based transfer learning method for leakage detection of pipeline under multiple working conditions with AE signals. Process Safety and Environmental Protection. 2023;170:1161-72.

[277] Zhou Q, Liu J, Liu L. Fast prediction of mine flow field based on convolution neural network. Process Safety and Environmental Protection. 2023;173:332-43.

[278] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)2016. p. 770-8.

[279] Peng L, Zhang J, Lu S, Li Y, Du G. One-dimensional residual convolutional

neural network and percussion-based method for pipeline leakage and water deposit detection. Process Safety and Environmental Protection. 2023;177:1142-53.

[280] Chen Y, Lin Y, Xu X, Ding J, Li C, Zeng Y, et al. Classification of lungs infected COVID-19 images based on inception-ResNet. Computer Methods and Programs in Biomedicine. 2022;225:107053.

[281] Mukherjee S, Huang X, Rathod VT, Udpa L, Deng Y. Defects Tracking via NDE Based Transfer Learning. 2020 IEEE International Conference on Prognostics and Health Management (ICPHM)2020. p. 1-8.

[282] Pan J, Gao L. A novel method for defects marking and classifying in MFL inspection of pipeline. International Journal of Pressure Vessels and Piping. 2023;202:104892.

[283] Li C, Yang F, Jia W, Liu C, Zeng J, Song S, et al. Pipelines reliability assessment considering corrosion-related failure modes and probability distributions characteristic using subset simulation. Process Safety and Environmental Protection. 2023;178:226-39.

[284] Zhou W, Xiang W, Hong HP. Sensitivity of system reliability of corroding pipelines to modeling of stochastic growth of corrosion defects. Reliability Engineering & System Safety. 2017;167:428-38.

[285] Zhu X. A comparative study of burst failure models for assessing remaining strength of corroded pipelines. Journal of Pipeline Science and Engineering. 2021;1:36-50.

[286] Ma B, Shuai J, Wang J, Han K. Analysis on the Latest Assessment Criteria of ASME B31G-2009 for the Remaining Strength of Corroded Pipelines. Journal of Failure Analysis and Prevention. 2011;11:666-71.

[287] Benjamin A, Andrade E. Modified Method for the Assessment of the Remaining Strength of Corroded Pipelines2003.

[288] Xu Y, Yan X, Sun B, Liu Z. Global contextual residual convolutional neural networks for motor fault diagnosis under variable-speed conditions. Reliability Engineering & System Safety. 2022;225:108618.

[289] Ye R, Li X, Ye Y, Zhang B. DynamicNet: A time-variant ODE network for multi-step wind speed prediction. Neural Networks. 2022;152:118-39.

[290] Zhou F, Li L, Zhang K, Trajcevski G. Urban flow prediction with spatial–temporal neural ODEs. Transportation Research Part C: Emerging Technologies. 2021;124:102912.

[291] Luo Z, Sun Z, Zhou W, Wu Z, Kamata S-i. Rethinking ResNets: improved stacking strategies with high-order schemes for image classification. Complex & Intelligent Systems. 2022;8:3395-407.

[292] Ricky, Rubanova Y, Bettencourt J, Duvenaud D. Neural Ordinary Differential Equations. arXiv pre-print server. 2019.

[293] Mehrafrooz B, Edalat P, Dyanati M. Cost consequence-based reliability analysis of bursting and buckling failure modes in subsea pipelines. Journal of Ocean Engineering and Science. 2019;4:64-76.

[294] De-León-Escobedo D. Risk-based maintenance time for oil and gas steel pipelines under corrosion including uncertainty on the corrosion rate and consequence-based target reliability. International Journal of Pressure Vessels and Piping. 2023;203:104927.

[295] Zhang S, Zhou W. Cost-based optimal maintenance decisions for corroding natural gas pipelines based on stochastic degradation models. Engineering Structures. 2014;74:74-85.

[296] Zhang H, Tian Z. Failure analysis of corroded high-strength pipeline subject to hydrogen damage based on FEM and GA-BP neural network. International Journal of Hydrogen Energy. 2022;47:4741-58.

[297] Li X, Jia R, Zhang R. A data-driven methodology for predicting residual strength of subsea pipeline with double corrosion defects. Ocean Engineering. 2023;279:114530.

[298] Zhang Y, Wang W, Shuai J, Shuai Y, Hua L, Lv Z. A novel assessment method to identifying the interaction between adjacent corrosion defects and its effect on the burst capacity of pipelines. Ocean Engineering. 2023;281:114842.

[299] Tee KF, Khan LR, Li H. Application of subset simulation in reliability estimation

of underground pipelines. Reliability Engineering & System Safety. 2014;130:125-31.
[300] Long Y, Huang S, Peng L, Wang S, Zhao W. A Characteristic Approximation Approach to Defect Opening Profile Recognition in Magnetic Flux Leakage Detection. IEEE Transactions on Instrumentation and Measurement. 2021;70:1-12.
[301] Piao G, Guo J, Hu T, Leung H, Deng Y. Fast reconstruction of 3-D defect profile from MFL signals using key physics-based parameters and SVM. NDT & E International. 2019;103:26-38.
[302] Kandroodi MR, Araabi BN, Ahmadabadi MN, Shirani F, Bassiri MM. Detection of natural gas pipeline defects using magnetic flux leakage measurements. 2013 21st Iranian Conference on Electrical Engineering (ICEE)2013. p. 1-6.
[303] Zhang M, Guo Y, Xie Q, Zhang Y, Wang D, Chen J. Estimation of Defect Size and Cross-Sectional Profile for the Oil and Gas Pipeline Using Visual Deep Transfer Learning Neural Network. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-13.
[304] Wang B, Guo Y, Wang D, Zhang Y, He R, Chen J. Prediction model of natural gas pipeline crack evolution based on optimized DCNN-LSTM. Mechanical Systems and Signal Processing. 2022;181:109557.
[305] Duan Z, Xie S, Huang L, Zhao R, Tian M, Liu T, et al. Quantitative sizing of compound location defects based on PECT-EMAT hybrid testing methods. Mechanical Systems and Signal Processing. 2022;178:109267.
[306] Eybpoosh M, Berges M, Noh HY. An energy-based sparse representation of ultrasonic guided-waves for online damage detection of pipelines under varying environmental and operational conditions. Mechanical Systems and Signal Processing. 2017;82:260-78.
[307] Chen J, Cao L, Song G. Detection of the pipeline elbow erosion by percussion and deep learning. Mechanical Systems and Signal Processing. 2023;200:110546.
[308] Ege Y, Coramik M. A new measurement system using magnetic flux leakage method in pipeline inspection. Measurement. 2018;123:163-74.
[309] Liu S, Sun Y, Jiang X, Kang Y. A new MFL imaging and quantitative

nondestructive evaluation method in wire rope defect detection. Mechanical Systems and Signal Processing. 2022;163:108156.

[310] Liu J, Ma Y, Zhang H, Su H, Xiao G. A modified fuzzy min–max neural network for data clustering and its application on pipeline internal inspection data. Neurocomputing. 2017;238:56-66.

[311] Yang L, Wang Z, Gao S, Shi M, Liu B. Magnetic flux leakage image classification method for pipeline weld based on optimized convolution kernel. Neurocomputing. 2019;365:229-38.

[312] Li F, Feng J, Lu S, Liu J, Yao Y. Convolution neural network for classification of magnetic flux leakage response segments. 2017 6th Data Driven Control and Learning Systems (DDCLS)2017. p. 152-5.

[313] Kim HM, Heo CG, Cho SH, Park GS. Determination Scheme for Accurate Defect Depth in Underground Pipeline Inspection by Using Magnetic Flux Leakage Sensors. IEEE Transactions on Magnetics. 2018;54:1-5.

[314] Sun H, Peng L, Huang S, Li S, Long Y, Wang S, et al. Development of a Physics-Informed Doubly Fed Cross-Residual Deep Neural Network for High-Precision Magnetic Flux Leakage Defect Size Estimation. IEEE Transactions on Industrial Informatics. 2022;18:1629-40.

[315] Jiang L, Zhang H, Liu J, Shen X, Wang L. Pipeline Irregular Defect Inversion for Magnetic Flux Leakage Detection System Based on Heterogeneous Multiclass Feature Fusion. IEEE Transactions on Instrumentation and Measurement. 2023;72:1-9.

[316] Bubenik T. 7 - Electromagnetic methods for detecting corrosion in underground pipelines: magnetic flux leakage (MFL). In: Orazem ME, editor. Underground Pipeline Corrosion: Woodhead Publishing; 2014. p. 215-26.

[317] Su S, Wang P, Shi P, Hao S, Liang T. Experiment and simulation on testing steel plate with corrosion defects via magnetic flux leakage method. Journal of Magnetism and Magnetic Materials. 2022;560:169595.

[318] Zhang H, Gao X, Unterman J, Arodz T. Approximation Capabilities of Neural ODEs and Invertible Residual Networks. In: Hal D, III, Aarti S, editors. Proceedings of

the 37th International Conference on Machine Learning. Proceedings of Machine Learning Research. PMLR; 2020. p. 11086--95.

[319] Nikola Kovachki ZL, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart , Anima Anandkumar. Neural Operator: Learning Maps Between Function Spaces. kovachki2023neural: arXiv; 2023.

[320] Zhang Q, Li J, Ding W, Ye Z, Meng Z. Mechanical fault intelligent diagnosis using attention-based dual-scale feature fusion capsule network. Measurement. 2023;207:112345.

[321] Pawan SJ, Rajan J. Capsule networks for image classification: A review. Neurocomputing. 2022;509:102-20.

[322] Zhu Z, Peng G, Chen Y, Gao H. A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis. Neurocomputing. 2019;323:62-75.

[323] Ru J, Lu B, Chen B, Shi J, Chen G, Wang M, et al. Attention guided neural ODE network for breast tumor segmentation in medical images. Computers in Biology and Medicine. 2023;159:106884.

[324] Cui W, Zhang H, Chu H, Hu P, Li Y. On robustness of neural ODEs image classifiers. Information Sciences. 2023;632:576-93.

[325] Lu J, Deng K, Zhang X, Liu G, Guan Y. Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens. iScience. 2021;24:102804.

[326] Zhang K, Tang B, Deng L, Tan Q, Yu H. A fault diagnosis method for wind turbines gearbox based on adaptive loss weighted meta-ResNet under noisy labels. Mechanical Systems and Signal Processing. 2021;161:107963.

[327] Yin J, Li J, Karimi IA, Wang X. Generalized reactor neural ODE for dynamic reaction process modeling with physical interpretability. Chemical Engineering Journal. 2023;452:139487.