Petroleum Science 23 (2026) 1487–1504

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

A hierarchical Bayesian–MCMC model for predicting the geometric dimensions of corrosion defects in wet gas gathering pipelines

Si-Jia Chen [a,*], Min Qin [a,**], Ke-Xi Liao [a], Yong-Chun Mu [b], Xiao-Dong Hao [b]

[a] Petroleum Engineering School, Southwest Petroleum University, Chengdu, 610500, Sichuan, China
[b] Petrochina Planning and Engineering Institute, Beijing, 100083, China

ARTICLE INFO

Article history:
Received 17 June 2025
Received in revised form
21 September 2025
Accepted 20 November 2025
Available online 26 November 2025

Edited by Teng Zhu

Keywords:
Corrosion defects
Defect dimensions
Hierarchical Bayesian
Markov Chain Monte Carlo
Corrosion prediction

ABSTRACT

Internal corrosion is a major threat to the safety of natural gas pipelines, with defect geometry—depth, length, and width—playing a critical role in structural integrity assessments. While corrosion depth prediction has been widely studied, systematic probabilistic modeling of defect length and width remains limited. This study develops a hierarchical Bayesian-Markov Chain Monte Carlo (HB-MCMC) framework to jointly predict corrosion defect dimensions from in-line inspection (ILI) data. The framework integrates non-centered parameterization and adaptive sampling to improve inference efficiency and employs a hierarchical dynamic thresholding procedure for robust data preprocessing and outlier filtering. Field data from two transmission pipelines in Southwest China, comprising 1845 defect records, are analyzed. Results demonstrate that defect length and width both increase with depth, with width exhibiting stronger sensitivity. Model diagnostics confirm convergence and reliable uncertainty quantification. To further explore underlying mechanisms, OLGA multiphase flow simulations are combined with statistical predictions, providing flow-parameter profiles along the pipelines and enabling correlation analysis between local hydrodynamics and defect geometry. The proposed framework not only enhances predictive capability for defect length and width but also provides new insights into flow-corrosion interactions under real operating conditions, offering a reproducible and data-driven tool for corrosion assessment.

1. Introduction

Pipelines serve as critical infrastructure for the long-distance transportation of oil and natural gas. However, statistics show that internal corrosion accounts for approximately 20%–40% of pipeline failures during long-term operation, with the proportion exceeding 50% in aging pipelines. Corrosion is a complex phenomenon influenced by multiple factors, including exposure time, environmental conditions, and material properties. The progression of corrosion defects is inherently stochastic, exhibiting variability across both time and space (Ma et al., 2023; El-Kady et al., 2024; Simonsen et al., 2025). Pipeline corrosion incidents not only lead to significant economic losses but also pose environmental hazards and risks to public safety. Therefore, accurately predicting the development of corrosion in pipelines is essential for implementing effective maintenance and repair strategies, thereby ensuring the safe operation of oil and gas transportation systems.

Currently, in Internal Corrosion Direct Assessment (ICDA), the remaining wall thickness is commonly used as a key indicator to evaluate the severity of corrosion, making the prediction of corrosion depth a central focus of research. Existing models for forecasting the growth of corrosion depth in oil and gas pipelines can be broadly categorized into three types: mechanistic models based on corrosion principles (Nesic et al., 1996, 2001; Papavinasam et al., 2006; Nesic, 2007, 2012; Qin, 2022), empirical formula models, and machine learning-based models (Seghier et al., 2020, 2022; Li et al., 2021, 2022; Hu, 2024; Huang et al., 2024; Song et al., 2025). Mechanistic models provide valuable physical insight but are often limited by parameter uncertainty. Empirical models are straightforward but lack transferability across operating conditions. Machine learning approaches show strong predictive ability, yet their black-box nature restricts

* Corresponding author.
** Corresponding author.
E-mail addresses: chensijia_42@163.com (S.-J. Chen), 202499010132@swpu.edu.cn (M. Qin).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2025.11.041

physical interpretability. However, according to widely used industry standards such as ASME B31G, RSTRENG (Effective Area Method), Shell 92 (Cosham et al., 2007), and DNV RP-F101, the three-dimensional characteristics of corrosion defects—including depth, length, and width—are also critical parameters for structural integrity assessments. While depth prediction has been extensively studied, systematic research on defect length and width prediction remains scarce.

Prior work has emphasized the role of defect geometry in structural integrity. Chen et al. (2021) investigated the impact of corrosion geometry on pipeline collapse pressure using finite element modeling, and found that complex corrosion shapes significantly compromise the collapse capacity of corroded pipelines. Tian et al. (2024) showed that corrosion depth, shape, and width directly affect the collapse resistance of subsea pipelines, while the influence of length diminishes with increasing size. These findings highlight that both length and width play decisive roles in residual strength assessment. Nevertheless, predictive models rarely capture the probabilistic distributions of these dimensions, especially their hierarchical dependence on depth.

Corrosion progression involves multiscale dynamic interactions among materials, environments, and fluid media, introducing considerable uncertainty (Qin et al., 2025; Sun et al., 2025; Yuan et al., 2025; Yue et al., 2025). Probabilistic frameworks, especially those based on Bayesian inference and Monte Carlo simulation, have been increasingly adopted to address this challenge. For instance, Wang et al. (2022) applied Bayesian–MCMC methods to model stochastic corrosion growth and failure probability. Martens et al. (2024) used Bayesian updating for crack width assessment. (Cui and Wang, 2024) integrated Bayesian frameworks with deep neural networks to capture nonlinear depth–factor relationships. Feng et al. (2025) combined Bayesian reasoning with a reverse diffusion model to improve depth prediction accuracy. Despite these advances, two key gaps remain in literature. First, research on the probabilistic prediction of defect length and width is still limited compared with the extensive work on depth. Second, existing approaches often fail to capture the hierarchical dependence between depth and the associated dimensional characteristics, leading to inadequate representation of defect morphology.

To address these issues, this study develops a posterior-propagation-based hierarchical Bayesian MCMC (HB-MCMC) framework for the probabilistic prediction of corrosion defect dimensions using in-line inspection (ILI) data. The model incorporates non-centered parameterization and adaptive sampling to enhance inference efficiency and employs a hierarchical dynamic thresholding method for robust data filtering. In addition, OLGA multiphase flow simulations are employed to provide flow parameter distributions, which are analyzed in conjunction with the statistical predictions to explore their influence on the spatial evolution of corrosion defects. This combined framework not only improves predictive capability for defect length and width but also provides new insights into flow-induced corrosion mechanisms under real operating conditions.

2. Data sources and preprocessing

This study is based on field-collected data from two natural gas transmission pipelines located in Southwest China. The primary data source is in-line inspection (ILI) records, providing a comprehensive dataset comprising 1845 corrosion defect records. All data was acquired using ILI tools and online monitoring systems. Each record includes information such as pipeline mileage, elevation, defect dimensions (length, width, and depth), clock orientation, and local wall thickness at the defect location. In this study, corrosion depth (expressed as a percentage) and wall thickness are used as input features, while defect length and width are treated as the prediction targets. Detailed pipeline specifications and operating conditions are summarized as follows.

2.1. Pipeline specifications and operating conditions

The fundamental specifications of the pipelines are summarized in Table 1. Both pipelines are made of L360Q low-carbon steel and adopt a variable wall thickness design. They have been in service for three years. The pipeline routes feature significant terrain undulations, with multiple low-lying segments prone to liquid accumulation.

The pipeline operating conditions are listed in Table 2. The transported natural gas contains acidic components (notably CO_2) that accelerate internal corrosion.

2.2. In-line inspection data and corrosion defect types

The in-line inspection (ILI) dataset is summarized in Table 3. Each record contains basic defect information (wall thickness and defect type), positional information (mileage, elevation, and clock orientation), and dimensional information (wall loss percentage, length, and width of the defect) (see Tables 4 and 5).

According to the statistical results (Fig. 1), a total of 1334 pitting corrosion defects were identified, accounting for 72.3% of all detected cases. Therefore, this study primarily focuses on pitting corrosion.

The classification of corrosion defect geometry (see Fig. 2) in this study follows the guidelines of the Pipeline Operators Forum (POF) (Forum, 2009), which define the dimensional parameters length L, width W, and depth d, and the parameter A. Specifically, $A = 10$ mm when $t < 10$ mm and $A = t$ when $t \geq 10$ mm, where t is the local nominal wall thickness.

2.3. OLGA multiphase flow simulation

To understand the flow conditions influencing corrosion along the pipelines, a transient multiphase flow model was developed using OLGA software (developed by Schlumberger). OLGA is an industry-standard tool for transient multiphase flow simulation, particularly suited for wet gas systems like those studied here. Based on the pipeline geometries (Section 2.1) and the inlet/outlet operating conditions (Table 2), a transient multiphase flow model was constructed for pipelines N-1 and N-2. This model simulated key flow parameters—such as temperature, pressure, flow regime, and liquid holdup—along the entire length of both pipelines (totaling 34.11 km, Figs. 3 and 4). The simulated parameter distributions were then used in conjunction with the spatially resolved ILI defect data to analyze potential correlations between local flow conditions and corrosion defect characteristics (see Fig. 5).

3. Machine learning methods

This section presents the probabilistic framework for predicting corrosion-defect dimensions from in-line inspection (ILI) data. We formulate a posterior-propagation hierarchical Bayesian model to jointly characterize defect depth, length, and width, adopt non-centered parameterization to reduce posterior dependence, and employ adaptive sampling to improve inference efficiency. A hierarchical dynamic thresholding procedure is used for robust data filtering prior to inference. The section is organized as follows: data preprocessing, model construction, and predictive framework.

Table 1
Basic specifications of the pipelines.

Pipeline	Outer diameter, mm	Wall thickness, mm	Service life, years	Distance, m	Elevation, m
N-1	219.1	8.8/10/11	3	15,816	220–346
N-2	508	17/20	3	17,540	220–316

Table 2
Key operating parameters.

Pipeline	DGP, 10^4 SCM/d	*P*, MPa		*T*, °C		Corrosion gas composition components, mol%	
		Inlet	Outlet	Inlet	Outlet	CO_2	O_2
N-1	60	8.9	6.4	21.0	5.0	5.371	0.009
N-2	627	7.9	6.0	17.0	7.1		

3.1. Data preprocessing

Corrosion defects with the same depth value may exhibit different lengths and widths, reflecting a one-to-many relationship. Statistically, both length and width tend to increase with greater defect depth. Based on this depth-dependent behavior, a hierarchical dynamic thresholding method is proposed in this study to preprocess the data.

A dynamic expansion factor (λ) is introduced, and the ILI data are discretized by corrosion depth into multiple levels (e.g., 0.1 mm, 0.2 mm, 0.3 mm, etc.). For each depth level, statistical metrics are calculated independently. The threshold is computed as shown in Eq. (1).

$$\text{Lower bound} = Q_1 - \lambda\cdot\text{IQR},\ \text{Upper bound} = Q_3 + \lambda\cdot\text{IQR} \quad (1)$$

Based on the calculated range of the hierarchical dynamic threshold, outlier filtering is performed using a two-dimensional joint labeling approach. If either length or width exceeds the threshold in any dimension, the data point is flagged as an outlier. The corrosion depth is mapped to a discrete index to preserve the hierarchical structure (e.g., 0.1 mm → 0, 0.2 mm → 1).

3.2. Model structure

3.2.1. Hierarchical Bayesian model

To quantify the impact of different corrosion depths on defect dimensions and their associated uncertainty, a hierarchical Bayesian model (HBM) is constructed. This model captures both the global trend and depth-specific effects through hierarchical parameterization.

The motivation for adopting a hierarchical Bayesian model lies in its ability to represent multi-level dependencies and borrow statistical strength across groups. In the context of corrosion data, defect dimensions are not only influenced by overall trends but also show systematic variations across different corrosion depth layers. Traditional regression models fail to capture such structured heterogeneity. The hierarchical structure allows us to model these depth-specific effects while sharing global information, which improves parameter estimation especially when sample sizes per group are limited.

Table 4
Example of dimensional combinations for pitting defects.

Depth, mm	Depth level	Number of samples	Length, mm	Width, mm
0.1	0	114	14	15
			18	16
			…	…
			22	29
0.2	1	52	17	26
			20	13
			…	…
			19	20
…	…	…	…	…

Table 5
Results of hierarchical dynamic threshold calculation.

Depth level, mm	Sample size	IQR		Threshold range, mm	
		Length	Width	Length	Width
0.1	114	7	9	(3.5, 31.5)	(2.5, 38.5)
0.2	52	6	10	(4, 28)	(3.2, 37)
0.3	44	4	3.5	(6, 22)	(4.12, 33.25)
…	…	…	…	…	…

The HBM is a statistical framework that introduces a layered structure to describe nested or grouped dependencies within the data. The formal definition is as follows:

$$\text{Observation level}:\ y_i \sim N\left(\alpha_{d[i]} + \beta\cdot d[i], \sigma^2\right) \quad (2)$$

$$\text{Depth}-\text{specific level}:\ \alpha_d \sim \left(\mu_{\text{global}}, \sigma^2_{\text{global}}\right),\ d = 1,\ 2,\ \ldots,\ D \quad (3)$$

$$\text{Global level}:\ \begin{cases} \mu_{\text{global}} \sim N\left(0, \tau_\mu^2\right) \\ \sigma_{\text{global}} \sim \text{HN}(\tau_\sigma) \end{cases} \quad (4)$$

Here, y_i denotes the standardized corrosion size, $d[i]$ represents the encoded corrosion depth, β is the global slope, σ is the standard deviation of observation noise, α_d denotes the intercept specific to each depth layer, D represents the total number of discretized depth layers, σ_{global} is the inter-level standard deviation, and μ_{global}, τ_μ and τ_σ are hyperparameters (typically set as fixed values).

The model consists of three hierarchical levels: the observation level, the depth-specific level, and the global level. At the

Table 3
In-line inspection data.

No.	Mileage, m	Elevation, m	O'clock orientation	Nominal wall thickness, mm	Type	Depth, %	Length, mm	Width, mm
1	1003.75	314.48	3:04	20.0	Pinhole	6.0	4	7
2	1116.38	304.74	11:48	17.0	Pitting	6.7	15	21
…	…	…	…	…	…	…	…	…
1844	1140.73	301.65	7:59	17.0	Pitting	5.2	26	29
1845	1150.61	301.09	9:25	17.0	Pitting	6.9	20	14

Defect type	Quantity	Proportion, %
Axial grooving	43	2.3
Axial slotting	86	4.7
Circumferential grooving	74	4.1
Circumferential slotting	101	5.5
General	24	1.3
Pinhole	183	9.9
Pitting	1334	72.3

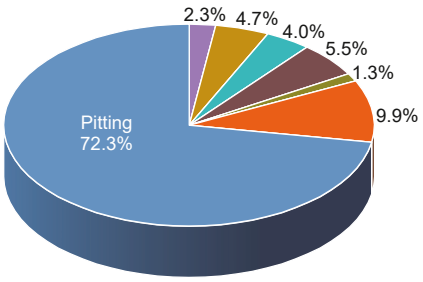


Fig. 1. Classification and proportion of corrosion defects.

Anomaly dimension class	Definition
General	$W \geqslant 3A$ and $L \geqslant 3A$
Pitting	$1A \leqslant W \leqslant 6A$ and $1A \leqslant L \leqslant 6A$ and $0.5 < L/W < 2$ and not $W, L \geqslant 3A$
Axial grooving	$1A \leqslant W \leqslant 3A$ and $L/W \geqslant 2$
Circumferential grooving	$1A \leqslant L \leqslant 3A$ and $L/W \leqslant 0.5$
Pinhole	$0 \leqslant W \leqslant 1A$ and $0 \leqslant L \leqslant 1A$
Axial slotting	$0 \leqslant W \leqslant 1A$ and $L \geqslant 1A$
Circumferential slotting	$0 \leqslant L \leqslant 1A$ and $W \geqslant 1A$

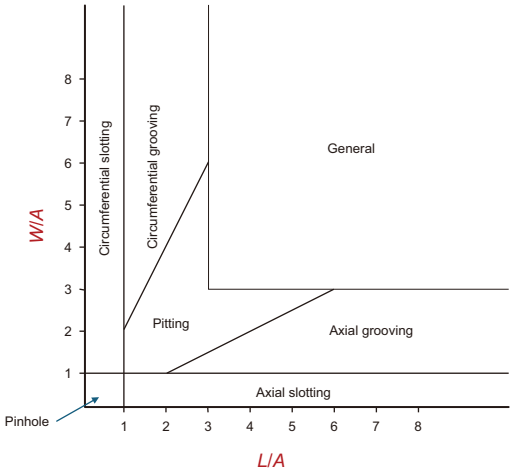


Fig. 2. Classification criteria for corrosion defect geometry based on POF.

Fig. 3. Pipeline N-1.

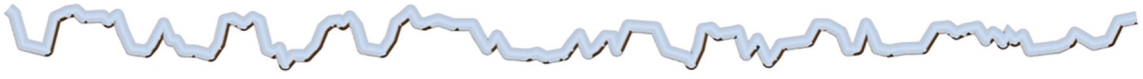


Fig. 4. Pipeline N-2.

observation level, defect dimensions (e.g., length or width) are modeled as a linear function of depth encoding and layer-specific intercepts. At the depth-specific level, the intercepts α_d vary across D layers and follow a global normal distribution. At the global level, the hyperparameters μ_{global}, σ_{global}, and β govern the prior distributions and induce parameter shrinkage across groups.

We assign weakly informative priors to these global parameters to preserve modeling flexibility while mitigating overfitting risks. These priors reflect general assumptions about parameter scales without imposing strict constraints, and are commonly used in hierarchical models to regularize group-level variation. This layered structure enables the model to simultaneously capture within-group variability and cross-group similarity, thereby improving robustness in parameter inference.

To address the low sampling efficiency caused by the strong correlation between α_d and σ_{global} in traditional centered hierarchical models, we adopt a non-centered reparameterization strategy. Specifically, an auxiliary variable ε_d is introduced to decompose the hierarchical effects into independent Gaussian variables, and a scaling factor γ is incorporated to prevent parameter overflow and improve sampling efficiency. This transformation enhances sampling efficiency and convergence behavior in the subsequent MCMC procedure. The reconstructed hierarchical intercept is expressed in Eq. (5). The overall structure of the hierarchical Bayesian model is illustrated in Fig. 6.

$$\alpha_d = \mu_{\text{global}} + \gamma \cdot \sigma_{\text{global}} \cdot \varepsilon_d,\ \varepsilon_d \sim N(0,1) \tag{5}$$

In essence, the model is organized into three layers: defect dimensions are first expressed at the observation level as functions of depth encoding with random noise; these are adjusted at the depth-specific level through intercepts that vary across depth groups; and finally, global parameters impose shared distributions at the top level. This explicit hierarchical structure ensures reproducibility by clearly defining parameter dependencies, while improving predictive robustness by pooling information across groups.

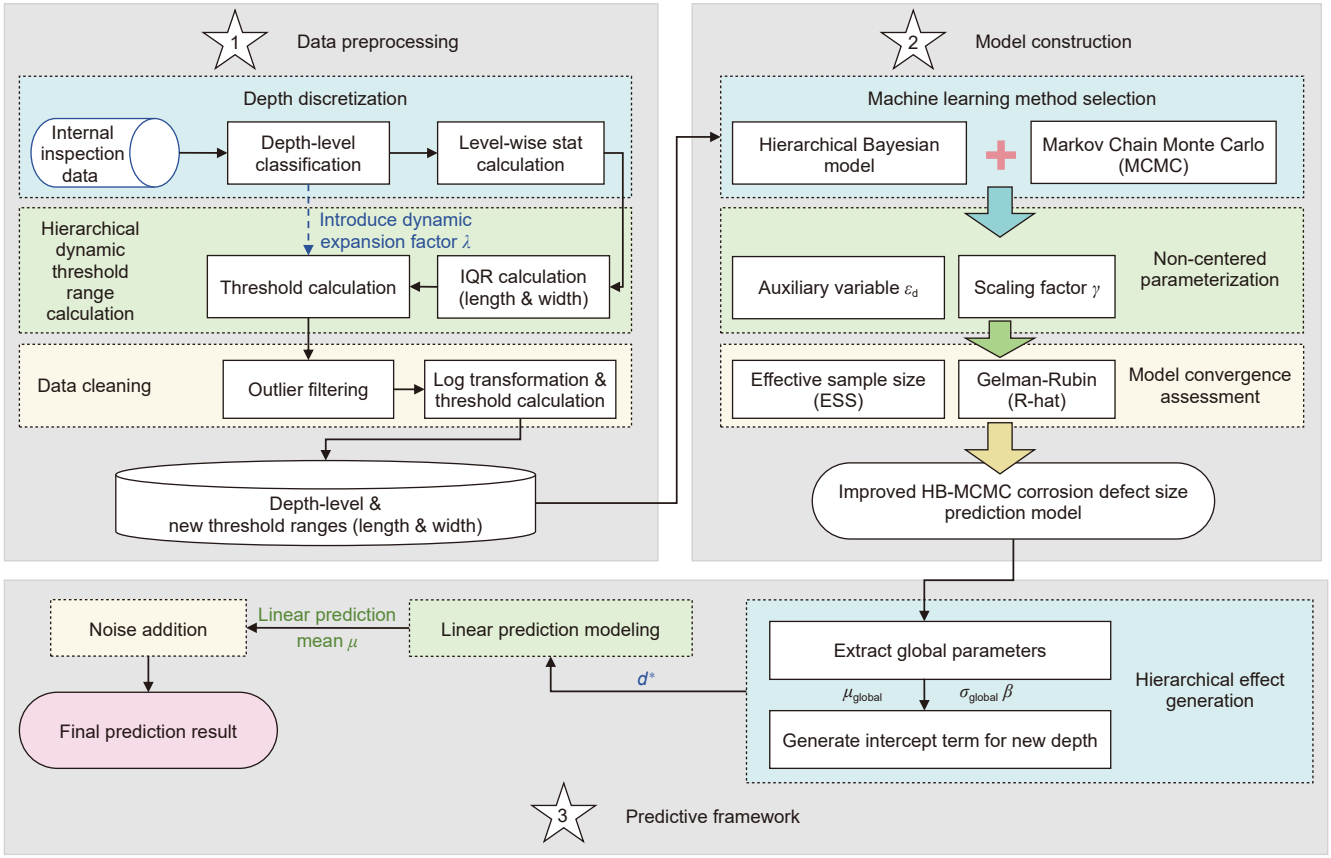


Fig. 5. Overall framework of corrosion defect size prediction model based on HB-MCMC.

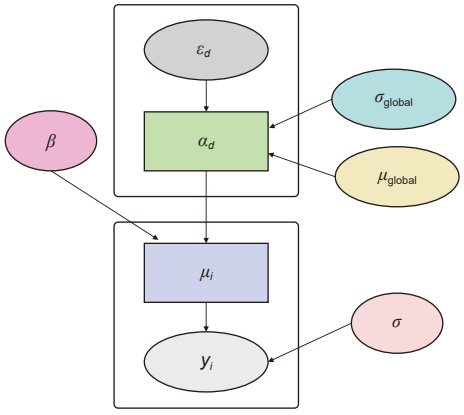


Fig. 6. Directed acyclic graph (DAG) of the Bayesian model structure.

3.2.2. Parameter inference with MCMC

To estimate the posterior distributions of parameters in the hierarchical Bayesian model described above, we employ Markov Chain Monte Carlo (MCMC) methods. Given the high-dimensional parameter space and non-conjugate posterior structure, MCMC provides an effective solution by generating samples from the target distribution.

In this study, we adopt the No-U-Turn Sampler (NUTS), an adaptive variant of the Hamiltonian Monte Carlo (HMC) algorithm. NUTS dynamically adjusts both the step size and path length during Hamiltonian dynamics simulation, automatically optimizing sampling efficiency and eliminating the need for manual tuning. This approach maintains detailed balance while significantly improving mixing performance for highly correlated parameters, such as hierarchical intercepts and variances, making it especially suitable for corrosion data with strong hierarchical dependencies.

Two independent Markov chains are configured for the prediction of length and width, respectively. Each chain includes 1000 burn-in iterations to eliminate the influence of initial values, followed by 1000 sampling iterations for posterior estimation. The target acceptance rate is set to 0.9 to enhance stability in high-curvature regions of the parameter space and reduce the risk of divergent transitions. The maximum tree depth is limited to 10 to balance computational efficiency with trajectory exploration capacity.

Model convergence was assessed using the Gelman–Rubin statistic (R-hat) and effective sample size (ESS). The R-hat statistic is computed as shown in Eq. (6); the model is considered to have converged when $\widehat{R}\to 1$.

$$\begin{cases} \widehat{R} = \sqrt{\dfrac{\widehat{Var}(\theta)}{W}} \\ \widehat{Var}(\theta) = \dfrac{N-1}{N}\cdot W + \dfrac{1}{N}\cdot B \end{cases} \tag{6}$$

Here, W denotes the within-chain variance, B denotes the between-chain variance, and N is the total number of samples per chain.

The effective sample size (ESS) is calculated as shown in Eq. (7):

$$\mathrm{ESS} = \frac{N}{1 + 2\sum_{k=1}^{\infty}\rho(k)} \tag{7}$$

Here, $\rho(k)$ denotes the autocorrelation coefficient at lag k and the summation is truncated at the first negative autocorrelation (Geyer, 1992). An ESS greater than 400 for all key parameters indicates that the sampling efficiency is sufficient to support reliable statistical inference.

Based on the posterior samples generated through MCMC, the statistical inference of key parameters is characterized by the posterior mean, standard deviation (SD), and highest density interval (HDI) (Hyndman, 1996). The posterior mean captures the central tendency of the parameter estimates, the posterior standard deviation quantifies the degree of uncertainty, and the HDI—one of the core indicators in Bayesian analysis—defines the narrowest interval that contains a specified probability mass (e.g., 95%) of the posterior distribution, thereby intuitively illustrating the plausible range of parameter values.

To further verify the adequacy of the non-centered parameterization and the efficiency of the NUTS sampling, visual diagnostics were conducted. For both the length and width models, trace plots were examined to confirm stable mixing and absence of non-stationary trends across chains, while autocorrelation plots verified the rapid decay of lag correlations.

3.3. Predictive framework

This study proposes a Bayesian predictive framework based on posterior propagation, which utilizes Monte Carlo sampling to generate predictive distributions and fully quantify prediction uncertainty. The prediction process consists of three stages: hierarchical effect generation, linear prediction modeling, and noise incorporation.

Global parameters—including the global mean μ_{global}, inter-level standard deviation σ_{global}, and global slope β—are extracted from the MCMC samples. Under the non-centered parameterization assumption, the intercept corresponding to a new corrosion depth d^* is generated as shown in Eq. (8). Then, the linear predictive mean μ is calculated using the depth encoding and global slope, as described in Eq. (9), capturing the trend of corrosion size variation with respect to depth. To account for observational noise and model residuals, the standard deviation σ is sampled from the posterior noise distribution, and the final predicted value y^* is generated accordingly. Notably, this sampling of σ does not rely on point estimates (e.g., posterior means); instead, posterior samples are drawn directly to retain uncertainty, ensuring that the prediction intervals maintain their nominal probability of covering the true values.

$$\alpha_{d^*} \sim N\left(\mu_{\text{global}}, \sigma^2_{\text{global}}\right) \tag{8}$$

$$\mu = \alpha_{d^*} + \beta \cdot d^* \tag{9}$$

$$y^* \sim N\left(\mu, \sigma^2\right) \tag{10}$$

4. Results

In this chapter, outlier detection is conducted to refine the defect dataset, followed by probabilistic prediction of corrosion-defect dimensions using the proposed Bayesian framework. The OLGA simulation is further employed to obtain flow-parameter distributions along the pipeline, enabling statistical comparison with defect characteristics.

4.1. Dynamic threshold calculation results

Based on the stratification of corrosion defect depths, different values of λ are selected to calculate the threshold ranges for length and width in each layer, as shown in Eq. (11).

A threshold of 1.2 mm is used to divide shallow and deep corrosion. This value falls within the typical 10%–20% wall loss range defined in pipeline integrity standards and is approximately balanced in sample count across the two groups, ensuring stable statistical inference. Given the variation in wall thickness across pipelines, absolute depth in millimeters was used for stratification instead of percentage-based criteria. This absolute-depth criterion inherently accounts for wall thickness differences, enabling the same threshold to be applied consistently to both pipelines without introducing bias. Although a fixed threshold of 1.2 mm was adopted in this study, the HB-MCMC framework itself is not restricted to this value and can be readily adapted to alternative stratification schemes when applied to pipelines with different wall thicknesses or operating conditions.

Accordingly, $\lambda = 1.5$ is applied for shallow corrosion to tolerate small surface variations, and $\lambda = 2.0$ for deep corrosion to enhance the detection of outlier geometries. The resulting thresholds are used for data cleaning and outlier filtering.

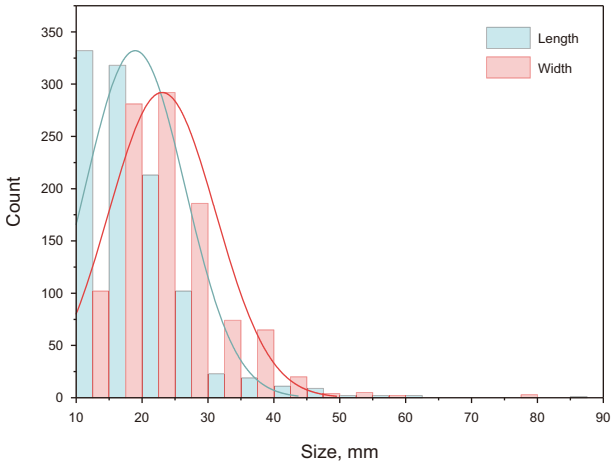
$$\lambda(d) = \begin{cases} 1.5 & d \le 1.2 \text{ mm (mild)} \\ 2.0 & d > 1.2 \text{ mm (severe)} \end{cases} \tag{11}$$

The cleaned size data still exhibited a right-skewed distribution, as illustrated in Fig. 7(a), with higher frequencies observed at smaller values of length and width. To stabilize the variance and eliminate scale differences, a logarithmic transformation followed by Z-score normalization was applied. The transformed data, shown in Fig. 7(c), approximately follows a standard normal distribution $N(0, 1)$ (see Fig. 8).

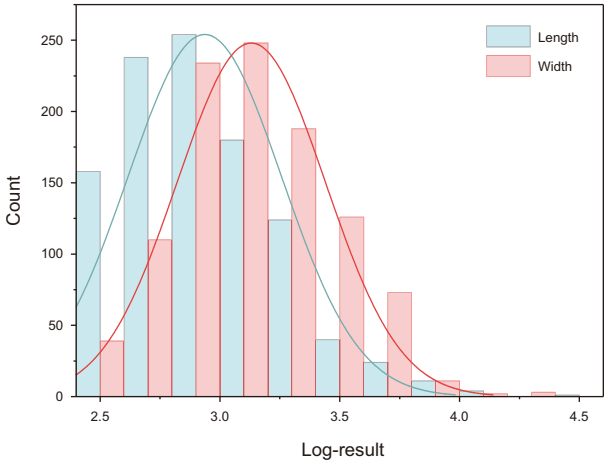
4.2. Model parameter calculation results

The HB-MCMC model was developed using the cleaned training dataset. To ensure robust algorithm performance, a set of hyperparameters was pre-defined before modeling, and auxiliary variable ε_d and scaling factor γ were introduced. Table 6 summarizes the posterior estimates and convergence diagnostics of key parameters for the length and width models. Based on a two-chain sampling design with 4000 iterations (2000 warm-up + 2000 sampling), all parameters yielded Gelman–Rubin statistics ($\hat{R}$) exactly equal to 1.00, and all effective sample sizes (ESS) exceeded 2000. This indicates that MCMC sampling achieved full convergence and sufficient efficiency. Sensitivity checks with shorter (500 warm-up + 500 sampling) and longer (2000 warm-up + 2000 sampling) chains produced consistent posterior estimates and convergence metrics, confirming that the chosen setting (1000 warm-up + 1000 sampling per chain) offers a good balance between computational efficiency and stability. Representative diagnostic plots for both models are provided in Appendix A (trace plots) and Appendix B (autocorrelation plots), which visually corroborate the numerical convergence indicators in Table 6.

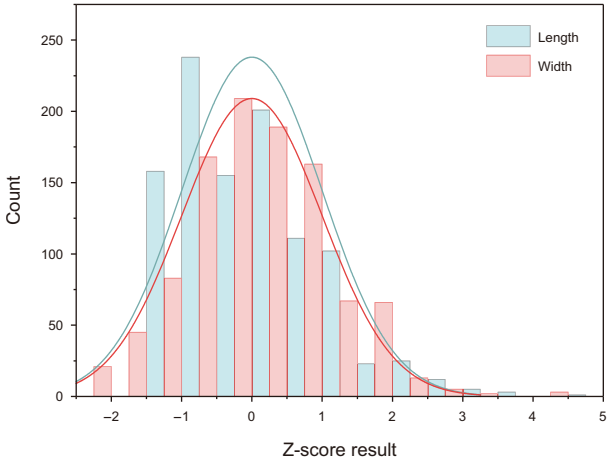
The global slope β represents the overall effect strength of defect depth on corrosion size (length and width). In the length model, β is estimated as 0.004, with a 95% highest density interval (HDI) of [0.003, 0.005], indicating that for every one-unit increase in standardized depth, the corrosion length increases by an average of 0.004 standardized units. In the width model, β is 0.007 with a 95% HDI of [0.005, 0.008], reflecting a 1.75 times stronger effect than in the length model and demonstrating a higher



(a) Original distribution of length and width



(b) Distribution after logarithmic transformation



(c) Distribution after Z-score normalization

Fig. 7. Distribution of length and width before and after transformation.

sensitivity to depth. In both models, the HDIs lie entirely within the positive domain, confirming the statistically significant positive association between depth and defect size. Given that pitting corrosion constitutes the majority of cases in the dataset, the reported depth sensitivity mainly reflects pitting behavior.

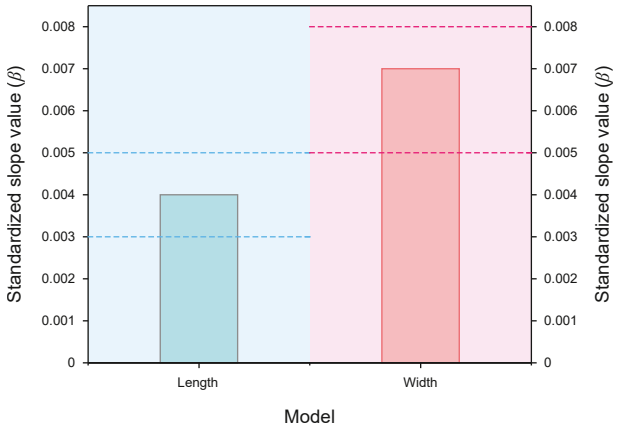


Fig. 8. Global effect of depth on corrosion dimensions.

The inter-level standard deviation σ_{global} captures the dispersion of corrosion baselines across different depth tiers, with smaller values indicating greater similarity among tiers. As illustrated in Table 6 and Fig. 9, the posterior mean of σ_{global} in the length model is 0.091 (95% HDI [0.000, 0.213]), with the peak density located at $\sigma_{global} = 0.033$ (normalized density: 6.28). In contrast, the width model exhibits a more compact posterior distribution, with a mean of 0.086 (95% HDI [0.000, 0.202]); its density peak shifts further left to $\sigma_{global} = 0.030$, where the highest density occurs at 0.003 (normalized density: 6.67). The Bhattacharyya coefficient between the two posterior distributions reaches 0.93, suggesting a high degree of similarity in both distribution shape and central tendency. These results confirm a consistent inter-level variation pattern across both models (see Fig. 10).

The observation noise parameter σ reflects the intrinsic uncertainty of the data and the residual error of the model. For the length model, σ is estimated at 0.960 with a 95% HDI of [0.921, 1], while for the width model, σ is 0.928 with a 95% HDI of [0.888, 0.967]. Both models exhibit HDI widths of 0.079 (<0.1), indicating high fitting accuracy. Notably, the noise level in the width model is 3.4% lower than that in the length model.

4.3. Improved HB-MCMC model development result

The improved HB-MCMC model was applied to predict the corrosion defect dimensions at unknown corrosion depths. Following the predictive framework, global parameters (μ_{global}, σ_{global}, and β) were first extracted from the posterior samples obtained from the dual MCMC chains. Based on the new depth d^*, the intercept was generated, and the linear predictive mean μ was calculated by combining the depth encoding with the global slope. Finally, the standard deviation σ was sampled from the posterior noise distribution to generate the final predicted value y^*.

For varying corrosion depths, the improved HB-MCMC model exhibits a significant positive correlation between defect size and depth (as shown in Table 7, Figs. 11 and 12). The predicted mean length of the defects increases progressively with depth: at a depth of 0.1 mm, the mean length is 16.66 mm; when the depth increases to 8.7 mm, the mean length reaches 21.19 mm. This trend is consistent with the physical mechanism of corrosion propagation (see Table 8).

Importantly, the uncertainty in the predicted length—measured by the width of the 95% highest density interval (HDI)—remains stable at 20–22 mm in the early stage (depth 0.1–0.5 mm), but expands to 27–28 mm at depths exceeding 8 mm.

Table 6
Posterior summaries of HB-MCMC model parameters.

Parameter	Mean		Std. Dev		95% HDI		ESS		$\widehat{R}$	
	Length	Width	Length	Width	Length	Width	Length	Width	Length	Width
μ_{global}	−0.248	−0.369	0.044	0.042	[−0.330, −0.164]	[−0.451, −0.293]	2974	2892	1	1
β	0.004	0.007	0.001	0.001	[0.003, 0.005]	[0.005, 0.008]	3197	2743	1	1
σ_{global}	0.091	0.086	0.070	0.063	[0, 0.213]	[0, 0.202]	2167	2005	1	1
σ	0.960	0.928	0.021	0.021	[0.921, 1]	[0.888, 0.967]	4579	4798	1	1

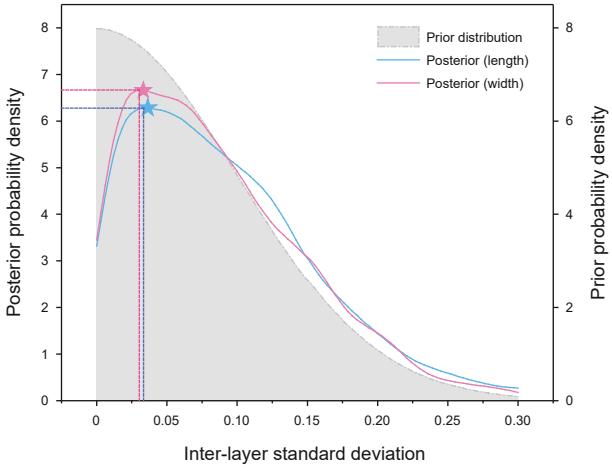


Fig. 9. Prior-posterior contrast of inter-layer standard deviation.

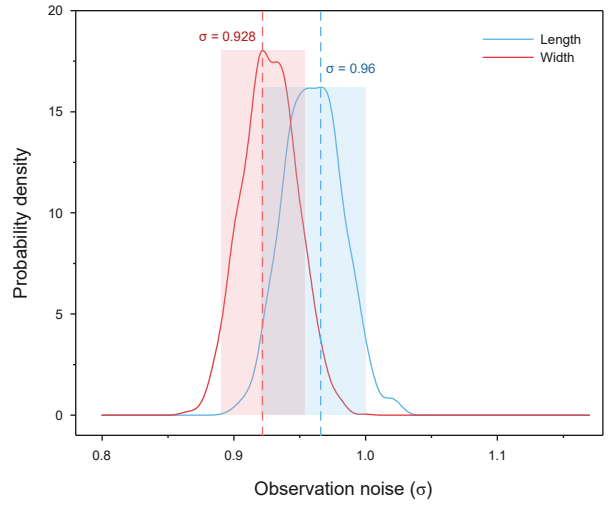


Fig. 10. Posterior distributions of observation noise parameters.

Table 7
Predicted corrosion defect dimensions.

Depth, mm	Length, mm			Width, mm		
	Mean	HDI_2.5	HDI_97.5	Mean	HDI_2.5	HDI_97.5
0.10	16.66	8.35	29.24	20.06	10.70	34.83
0.11	16.95	8.68	29.66	19.92	10.85	33.32
0.12	17.15	8.66	30.49	20.18	10.66	34.24
0.13	17.11	8.70	29.91	20.16	11.08	34.80
0.14	16.95	8.63	29.47	19.86	10.94	33.66
…	…	…	…	…	…	…

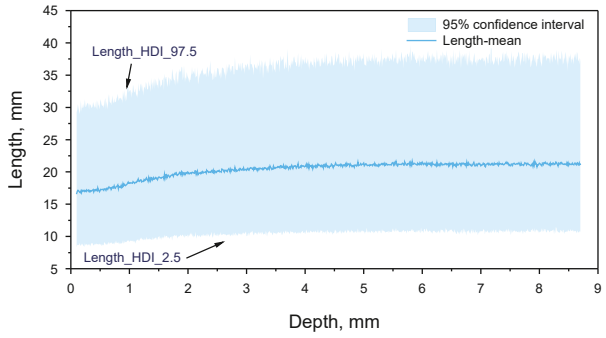


Fig. 11. Predicted trend of pitting length with corrosion depth.

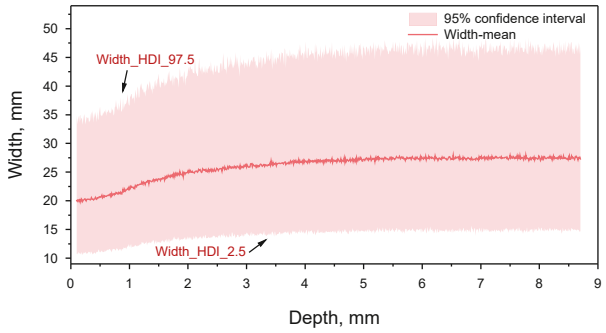


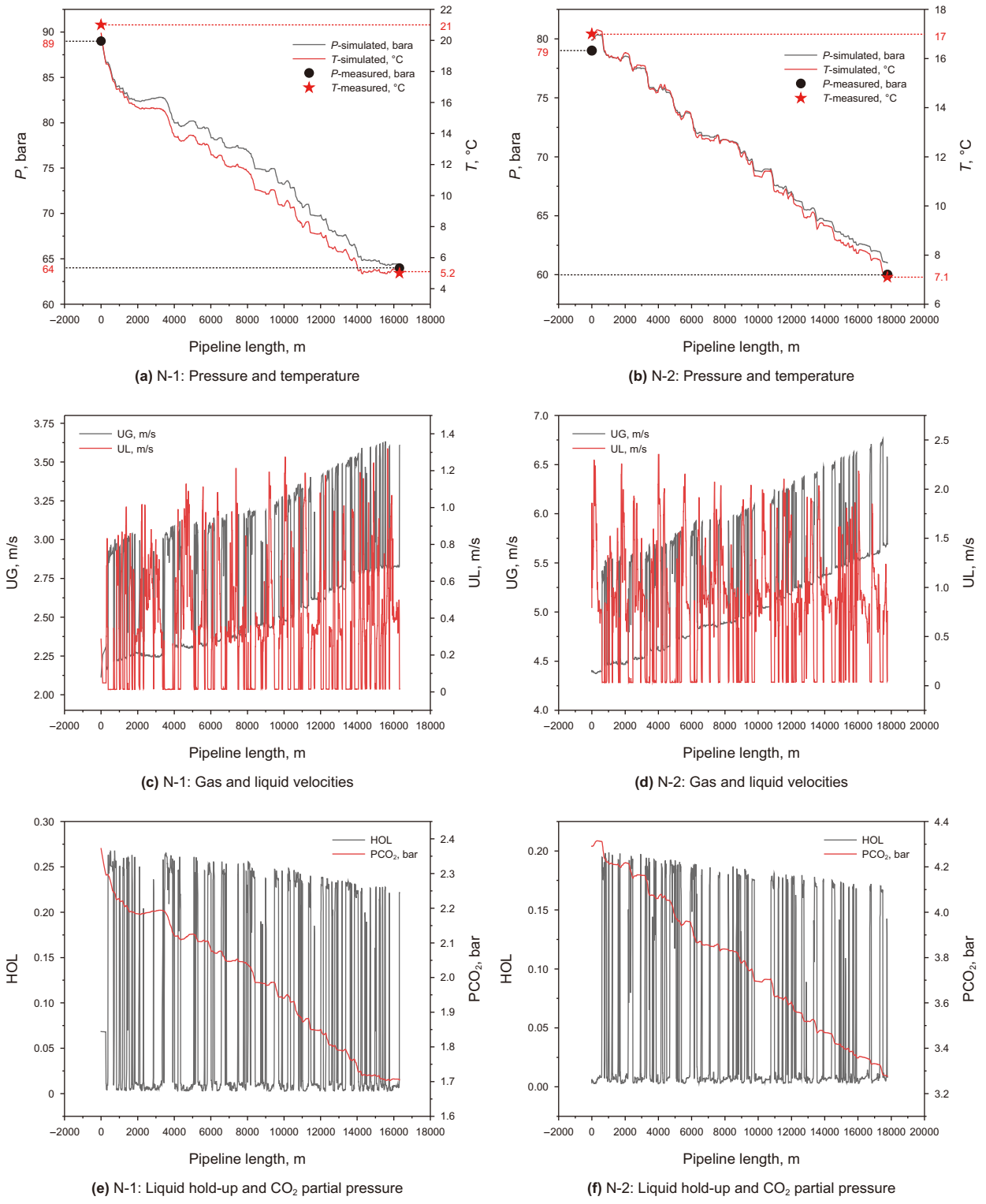
Fig. 12. Predicted trend of pitting width with corrosion depth.

Table 8
Summary of pitting dimension outliers by terrain type.

Terrain type	Number of length outliers	Number of width outliers
Uphill	10	1
Downhill	10	1
low-lying	0	13
Horizontal	1	0
Total	**21**	**15**

This indicates that multi-factor coupling effects become more prominent in regions with severe corrosion. Compared with length, the predicted width demonstrates stronger sensitivity to depth, with a 36% increase in the mean value (from 20.06 mm to 27.29 mm) and a 25% expansion in the HDI width (from 24 mm to 30 mm).

The prediction results reveal that the mean width of corrosion defects consistently exceeds the mean length, indicating a characteristic “wider-than-longer” geometry of pitting in natural gas pipelines. Moreover, the HDI interval widths for length and width exhibit strong synchronicity (Pearson $r = 0.83$, $p < 0.01$), suggesting that the uncertainty in size prediction primarily arises from shared environmental fluctuation factors—such as fluid



(a) N-1: Pressure and temperature

(b) N-2: Pressure and temperature

(c) N-1: Gas and liquid velocities

(d) N-2: Gas and liquid velocities

(e) N-1: Liquid hold-up and CO_2 partial pressure

(f) N-2: Liquid hold-up and CO_2 partial pressure

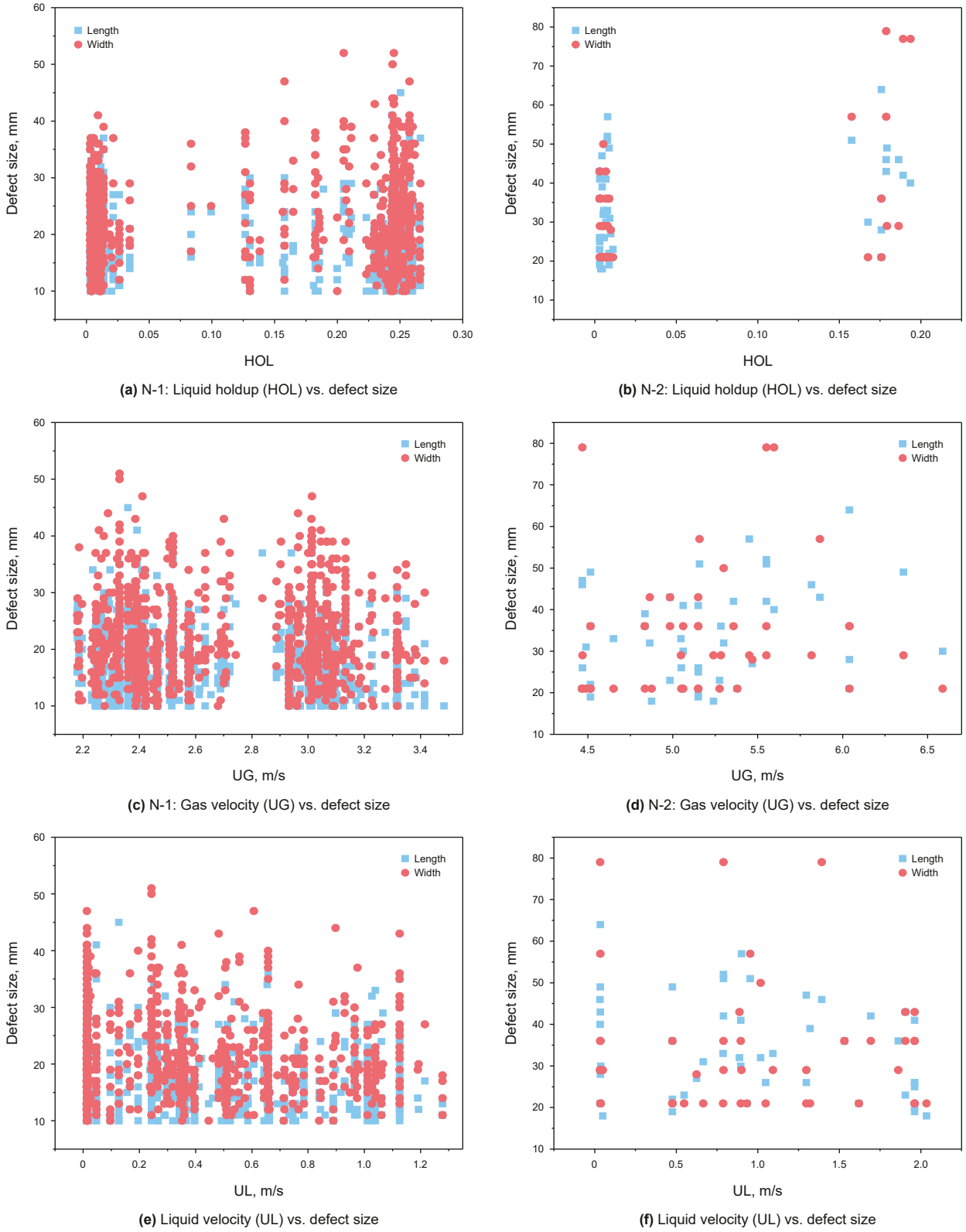
Fig. 13. Spatial distribution of multiphase flow parameters along pipelines N-1 and N-2.

composition, external conditions, and operational states—rather than from independent measurement noise.

In this study, length and width were selected as the target geometric parameters because they can be obtained from ILI data with relatively high confidence and are directly used in widely adopted pipeline integrity assessment models (e.g., ASME B31G, DNV RP-F101). Other descriptors such as defect area or aspect ratio, while potentially valuable, were excluded due to the current dataset's limited consistency and will be considered in future work.

4.4. Distribution of flow parameters along the pipeline

Two gas transmission pipelines from the same gas field, designated N-1 and N-2, with a total length of 34.11 km, were



(a) N-1: Liquid holdup (HOL) vs. defect size

(b) N-2: Liquid holdup (HOL) vs. defect size

(c) N-1: Gas velocity (UG) vs. defect size

(d) N-2: Gas velocity (UG) vs. defect size

(e) Liquid velocity (UL) vs. defect size

(f) Liquid velocity (UL) vs. defect size

Fig. 14. Scatter plots of defect size versus key flow parameters for pipelines N-1 and N-2.

modeled using the OLGA multiphase flow simulator. The spatial distributions of temperature, pressure, velocity, and CO_2 partial pressure were simulated (see Fig. 13). The deviations between the simulated and measured operating pressures and temperatures (both inlet and outlet) are within the range of 0.47%–4%, which falls within acceptable engineering tolerance. Due to terrain undulations and ambient temperature variations, both pipelines exhibit similar oscillatory trends in pressure and temperature

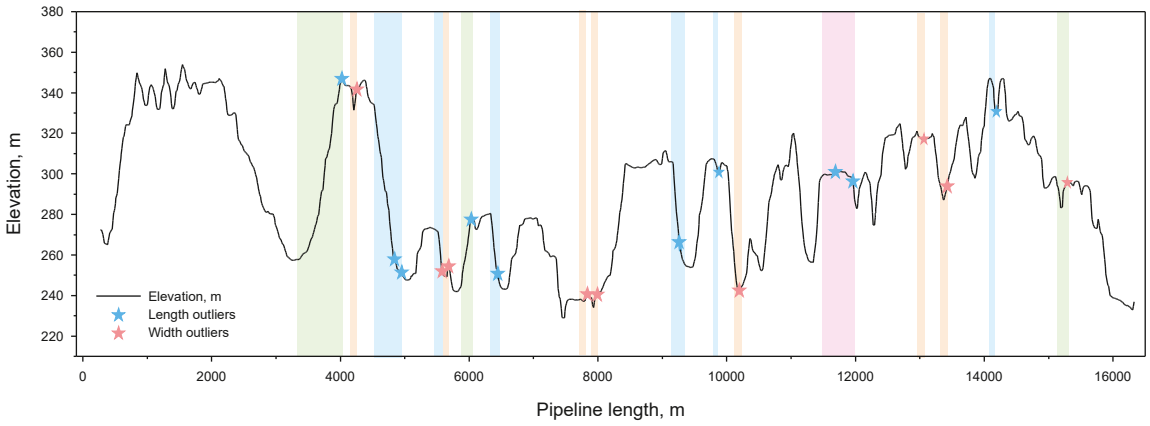


Fig. 15. Spatial distribution of pitting size outliers in pipeline N-1.

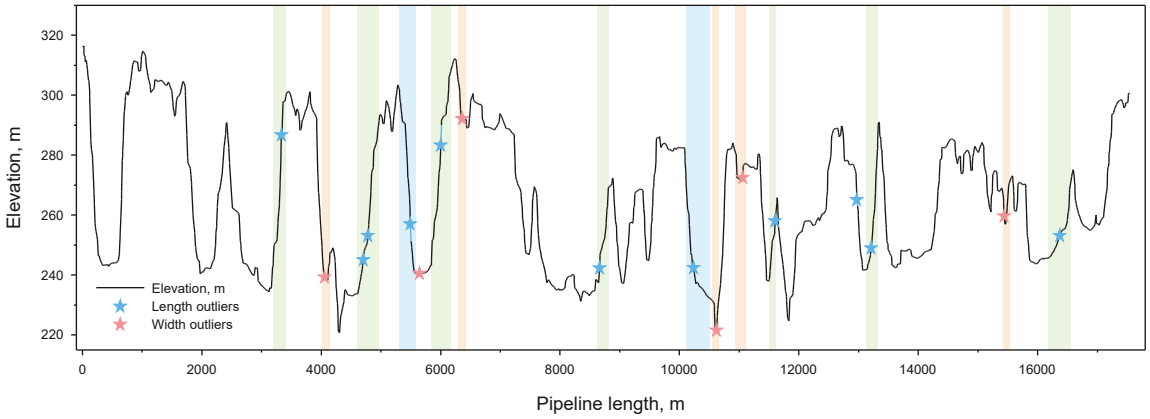


Fig. 16. Spatial distribution of pitting size outliers in pipeline N-2.

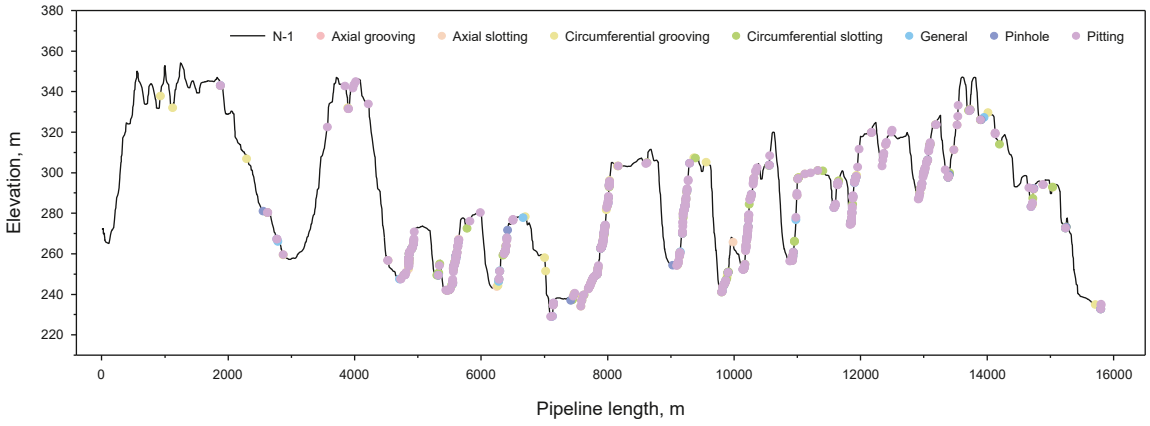


Fig. 17. Longitudinal distribution of corrosion defects in pipeline N-1.

drops. In pipeline N-1, the pressure decreases from 89.5 bara to 64.3 bara ($\Delta P = 28.0\%$), and the temperature drops from 20.5 °C to 5.2 °C ($\Delta T = 15.3$ °C). In pipeline N-2, the pressure declines from 79.9 bara to 61.0 bara ($\Delta P = 23.0\%$), while the temperature falls from 16.8 °C to 7.2 °C ($\Delta T = 9.6$ °C). Along with the pressure decay, the CO_2 partial pressure in both pipelines shows a monotonic decreasing trend.

Compared to pipeline N-1, pipeline N-2 exhibits a higher gas velocity (4.36–6.38 m/s). The enhanced gas-phase carrying capacity in N-2 effectively suppresses liquid accumulation, resulting in a 25.8% lower hold-up rate (HOL = 0.0027–0.1979) compared to N-1 (HOL = 0.003–0.267). Spatial distribution analysis reveals that liquid hold-up peaks (HOL > 0.2) tend to occur in terrain depressions and uphill segments due to reduced local flow velocities. Conversely, downhill segments, driven by gravitational acceleration, show elevated liquid velocities, causing HOL to approach zero. It should be noted, however, that the validation of the OLGA simulations was limited to inlet and outlet operating

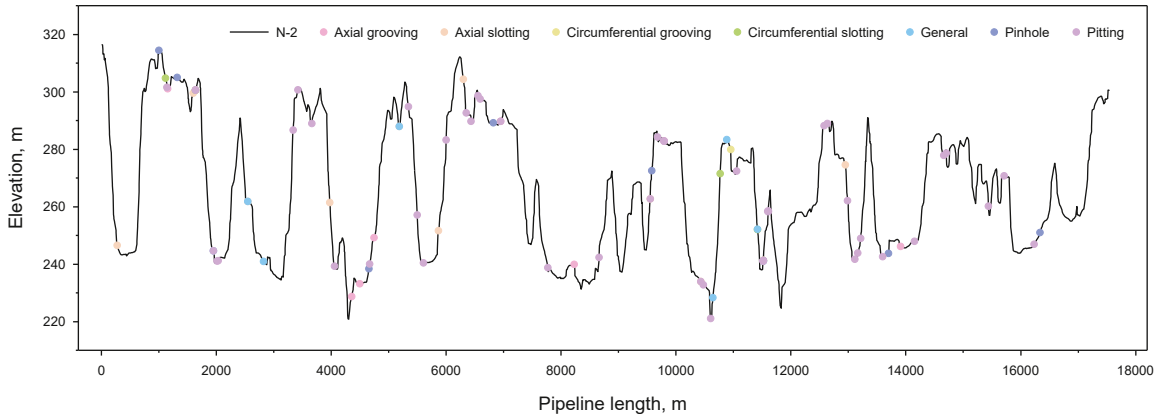


Fig. 18. Longitudinal distribution of corrosion defects in pipeline N-2.

Table 9
Mean values and ranges of pitting defect dimensions at different pipeline locations.

Pipeline segment	Parameter		N-1	N-2
Uphill	Length, mm	Mean	17.55	34.72
		Range	11–45	18–64
	Width, mm	Mean	22.29	29.5
		Range	11–52	21–79
Downhill	Length, mm	Mean	18.58	33.9
		Range	11–34	18–52
	Width, mm	Mean	23.06	37.6
		Range	11–47	21–79
Horizontal	Length, mm	Mean	19.09	–
		Range	11–37	–
	Width, mm	Mean	22.46	–
		Range	11–52	–
Low-lying	Length, mm	Mean	23.4	38.25
		Range	20–35	32–46
	Width, mm	Mean	30.7	46.5
		Range	22–50	36–79

conditions, as field-measured internal flow profiles are not available in practice. Consequently, the spatial distributions of flow parameters reported here should be interpreted as simulation-based estimates rather than direct measurements.

To further explore the potential influence of flow parameters on pitting defect dimensions, defect length and width data extracted from in-line inspection results were plotted against the simulated flow parameters liquid holdup (HOL) gas velocity (UG), liquid velocity (UL) at corresponding spatial locations. Scatter plots for both pipelines are shown in Fig. 14. No clear monotonic or linear trends are observed in these plots, indicating that defect dimensions are not strongly dominated by any single flow parameter. Instead, the results highlight that pitting defect growth is influenced by the coupled effects of flow hydrodynamics and terrain, rather than a single dominant factor (see Fig. 15).

4.5. Analysis of outliers in pitting corrosion defect dimensions

Building on the scatter-plot findings in Section 4.4, which indicate multi-parameter interactions rather than a single dominant driver, we further examined extreme cases through outlier analysis. This approach highlights defect dimensions that deviate significantly from the expected ranges and allows their spatial distribution to be interpreted in relation to terrain and flow conditions.

Based on the dimensional thresholding results presented in Section 4.1, outlier detection was performed on pitting corrosion defects in the two wet gas pipelines. A total of 36 outliers were identified, including 21 length-related anomalies (marked in blue) and 15 width-related anomalies (marked in red). As illustrated in the figure, green regions represent uphill pipeline segments, yellow regions correspond to terrain depressions, blue regions indicate downhill sections, and pink areas denote horizontal segments.

The anomaly determination for both pipelines is based on defect dimensions exceeding the threshold range of length and width under corresponding depth conditions. Compared to defects within the normal range, these anomalous defects exhibit significantly larger sizes and show a stronger tendency to evolve into extensive corrosion areas. Combined with elevation data, a clear spatial correlation is observed between the distribution of anomalies and the terrain features along the pipeline route.

(1) Distribution characteristics of length anomalies

Length anomalies are predominantly concentrated in regions with significant terrain variations—namely, uphill and downhill segments. Taking the uphill section at 3332.4 m of Pipeline N-2 as an example, the gas velocity (UG) is 5.42 m/s (high), the liquid holdup (HOL) is 0.19 (moderate), the liquid velocity (UL) is 0.03 m/s (indicating slow movement due to gas-phase drag), and the wall gas-phase shear stress (TAUWG) reaches 6.50×10^{-5} bara (a high-shear zone). In such conditions, the gas–liquid mixture flows rapidly along the axial direction, producing axial scouring on the pipe wall. Due to liquid retention on the uphill segment and fluctuations in liquid velocity, entrained droplets impact the wall and generate linear erosion marks, facilitating axial defect propagation in the flow direction.

In the downhill section at 5592.5 m of Pipeline N-2, UG is 4.73 m/s (high), UL is 2.29 m/s (high), and HOL is 0.003 (low). Here, gravity acts in concert with the flow direction, accelerating the liquid phase and substantially increasing UL, while liquid retention is significantly reduced. This leads to a sharp drop in HOL and a "thin and fast" liquid film regime. The shear stress from the liquid on the wall becomes uniformly distributed axially, while the gas-phase shear stress remains high, accelerating axial growth of corrosion defects.

A small number of length anomalies are found in horizontal segments. At 11,638.2 m of Pipeline N-1, UG is 2.61 m/s (moderate), UL is 0.35 m/s (low), and HOL is 0.12 (moderate). Although the flow velocities exhibit only slight fluctuations, prolonged stable flow conditions can continuously erode corrosion products on the pipe wall, gradually stripping them away and promoting the axial development of corrosion defects.

(2) Distribution characteristics of width anomalies

Width anomalies tend to occur in low-lying segments of the pipeline. For example, at 4217.2 m of Pipeline N-1, the liquid holdup (HOL) reaches 0.26 (high), the liquid velocity (UL) is 0.23 m/s (indicating liquid retention), the gas velocity (UG) is 2.92 m/s (moderate), and the wall shear stress from the liquid phase (TAUWWT) is 0.0002 bar—significantly higher than in low-HOL regions. In this scenario, a persistent liquid film covers the bottom of the pipe wall, facilitating the circumferential expansion of corrosion depth. In low-lying areas, liquid retention leads to a substantial reduction in liquid velocity, with flow occurring predominantly in laminar or stagnant regimes. This stable liquid film continuously wets the bottom of the pipe, promoting CO_2 dissolution and enhancing circumferential wall shear stress from the liquid phase, thereby accelerating the circumferential development of bottom-wall corrosion defects.

5. Discussion

This chapter provides a comprehensive discussion of the results presented in Chapter 4, focusing on the implications of defect distribution, dimensional evolution, and their correlations with flow parameters, as well as mechanistic interpretations of the observed patterns (see Fig. 16).

5.1. Correlation between defect distribution and flow parameters

Based on in-line inspection data, multiphase flow simulations, and the predictive model, the evolution of corrosion defect dimensions in the wet gas pipelines was analyzed. The longitudinal distributions of corrosion defects along pipelines N-1 and N-2 are shown in Figs. 17 and 18.

Overall, pipeline N-1 exhibits a significantly higher number of corrosion defects than pipeline N-2, with defects predominantly concentrated in uphill sections and low-lying terrain. In contrast, pipeline N-2 shows relatively fewer defects, mainly located in undulating and low-lying areas. Although both pipelines transport the same gas composition, pipeline N-2 handles approximately ten times the throughput of N-1, leading to distinct differences in internal flow characteristics and corrosion behavior.

As discussed in Section 4.4, OLGA simulation results indicate that gas velocity in pipeline N-2 (4.36–6.76 m/s) is substantially higher than that in N-1 (0.70–1.26 m/s). The stronger gas-phase shear force in N-2 enhances liquid entrainment and significantly reduces the likelihood of liquid holdup, effectively limiting the accumulation of corrosive media. This gas-dominant flow regime accounts for the relatively lower incidence of corrosion defects observed in pipeline N-2.

5.2. Correlation between defect dimensions and flow parameters

According to Table 9 the corrosion defects in both wet gas pipelines generally exhibit a “width greater than length” characteristic across all segments—except for the uphill section of pipeline N-2. This suggests that corrosion tends to propagate circumferentially rather than longitudinally along the flow direction. The mean values and ranges of defect sizes in different sections of pipeline N-1 are consistently smaller than those in pipeline N-2, indicating that although corrosion occurs more frequently in N-1, it is typically confined to a smaller scale.

A comparative analysis of flow conditions and corrosion characteristics across different terrain segments further elucidates the relationship between defect geometry evolution and localized flow parameters, as detailed below.

(1) Uphill segment

In the uphill sections, both pipelines exhibit localized peaks in liquid holdup due to the difficulty of liquid “climbing” against gravity. However, pipeline N-2 maintains a stable gas-dominated flow regime characterized by a relatively high gas velocity (mean: 5.54 m/s) and low liquid velocity (0.38 m/s). This flow pattern reduces the duration of liquid film coverage while enhancing axial shear disturbance, which promotes the elongation of corrosion pits along the flow direction—resulting in a geometric pattern where length exceeds width. In contrast, pipeline N-1 exhibits a much lower gas velocity (1.1 m/s) and higher liquid velocity (0.74 m/s), leading to significant liquid accumulation. Consequently, this section displays a high density of corrosion defects with large variability in size.

(2) Downhill segment

In the downhill sections, gravitational acceleration facilitates the formation of high-speed thin liquid films and increases liquid shear velocity. Although this flow regime helps reduce film thickness and liquid retention, the elevated shear stress may intensify localized wall erosion. This explains the observation of fewer but slightly larger defects, a trend particularly evident in the downhill section of pipeline N-1.

(3) Horizontal segments

In horizontal pipeline sections, the influence of gravity is minimal, resulting in relatively stable flow structures and low-frequency gas-liquid interface disturbances. Consequently, the local corrosion environment exhibits limited variability. Corrosion defects in these segments tend to be uniformly distributed, with moderate sizes and low fluctuation in geometry.

(4) Low-lying areas

Low-lying segments are the most prone to liquid accumulation due to the combined effects of terrain depressions and reduced flow inertia. These conditions significantly increase the local liquid holdup. Both N-1 and N-2 pipelines exhibit a concentration of corrosion defects with relatively larger dimensions in these regions. Compared to other segments, the corrosion pits in low-lying areas show a more pronounced “width greater than length” characteristic. Statistical analysis reveals that the average length-to-width ratios of pitting defects in the low-lying sections are 0.76 (N-1) and 0.82 (N-2), the lowest among all segments. This indicates that corrosion in these areas predominantly propagates radially. The extended liquid residence time and low flow velocity, coupled with a lack of sufficient axial shear disturbance, facilitate uniform corrosion propagation across the local surface. This leads to the formation of wide and shallow dish-shaped pits, posing a higher threat to structural integrity.

5.3. Multi-source uncertainties in corrosion defect prediction

In the present work, defect-size predictions were developed by combining ILI measurements with OLGA-simulated flow parameters within a hierarchical Bayesian framework that quantifies parameter uncertainty via posterior sampling. Nonetheless, it is important to recognize that predictive performance in real pipeline integrity management is further affected by multi-source uncertainties that extend beyond the sampling variability captured by the HB-MCMC. These uncertainties include, but are not limited to, measurement errors and potential systematic biases in ILI sizing,

modeling assumptions and input uncertainties in multiphase flow simulations, and dependencies that may exist across information sources and over time. Such effects can interact and accumulate, potentially biasing point estimates and altering predictive intervals in ways not reflected by single-source treatments (Der Kiureghian and Ditlevsen, 2009; Ali et al., 2022; Qin, 2022). Future work should therefore explicitly model measurement bias, spatiotemporal dependency of inspection data, and uncertainty propagation from simulation inputs to predictive outputs, to improve the robustness of prediction and decision support.

6. Conclusions

This study proposes a systematic Bayesian prediction framework based on posterior propagation for estimating corrosion defect sizes in natural gas pipelines, integrating mechanistic analysis, Bayesian inference, and operational data. The main conclusions are as follows.

(1) In gas transmission pipelines, pitting defect size shows a significant positive correlation with depth. The sensitivity of width to depth is approximately 1.75 times that of length, indicating a geometric tendency of "width greater than length".
(2) The hierarchical Bayesian MCMC (HB-MCMC) model successfully quantifies predictive uncertainty through posterior sampling. The 95% highest density intervals (HDIs) for predicted length and width span 20–28 mm and 24–30 mm, respectively. The hierarchical structure preserves depth-specific corrosion kinetics while reducing inter-layer variability.
(3) Based on OLGA multiphase flow simulation, defect points exceeding the length threshold tend to occur in regions with significant terrain variation (i.e., uphill and downhill segments), with a few found in horizontal segments with relatively stable flow. In contrast, width outliers are predominantly located at low-lying segments where liquid accumulation occurs.
(4) The HB-MCMC framework can be integrated with real-time ILI data updates by sequentially incorporating new measurements into the posterior distribution, enabling timely refinement of defect size predictions. Although developed and validated using ILI data from two wet gas pipelines in southwestern China, the hierarchical structure is adaptable to datasets from other regions and operating conditions. This contribution advances pipeline integrity assessment by enabling probabilistic prediction of full defect morphology rather than depth alone, while future work will extend validation to pipelines with different fluid compositions, materials, and environmental conditions to enhance generalizability.

CRediT authorship contribution statement

Si-Jia Chen: Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Min Qin:** Writing – review & editing, Investigation, Funding acquisition, Formal analysis. **Ke-Xi Liao:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Yong-Chun Mu:** Validation, Project administration, Formal analysis. **Xiao-Dong Hao:** Visualization, Project administration, Methodology, Investigation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (52174062), Sichuan Youth Fund Project (2025ZNSFSC1366) and China Postdoctoral Science Foundation (2025M772957).

Appendix A. Trace plots for key parameters

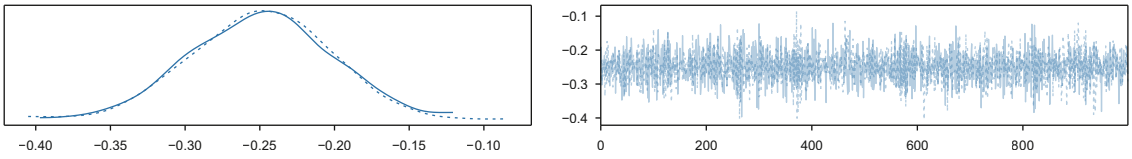


Fig. A1. Trace plots—length model (μ_{global}).

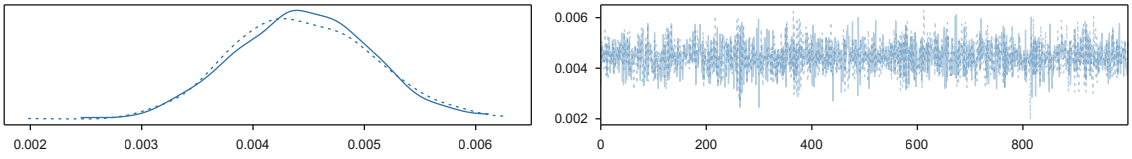


Fig. A2. Trace plots—length model (β).

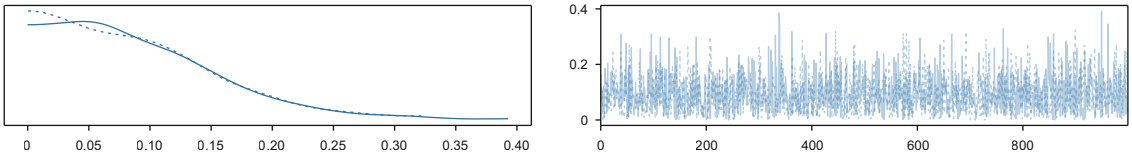


Fig. A3. Trace plots—length model (σ_{global}).

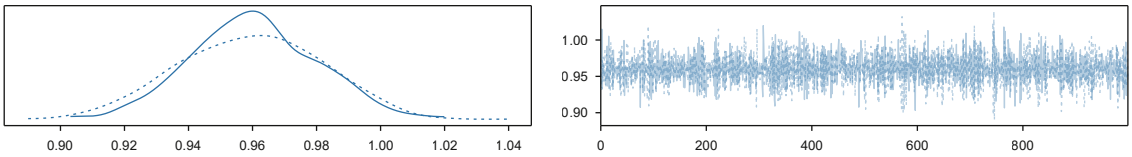


Fig. A4. Trace plots—length model (σ).

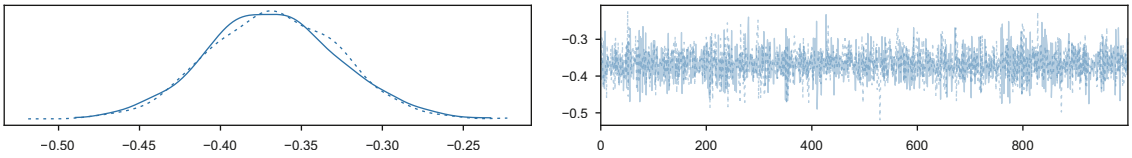


Fig. A5. Trace plots—width model (μ_{global}).

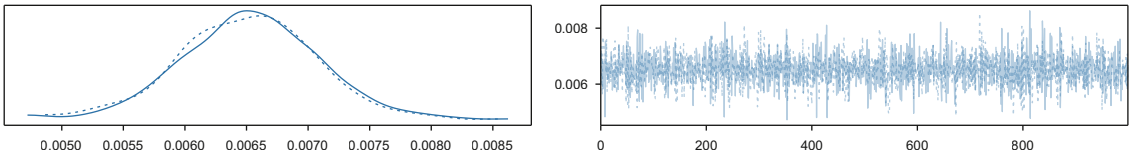


Fig. A6. Trace plots—width model (β).

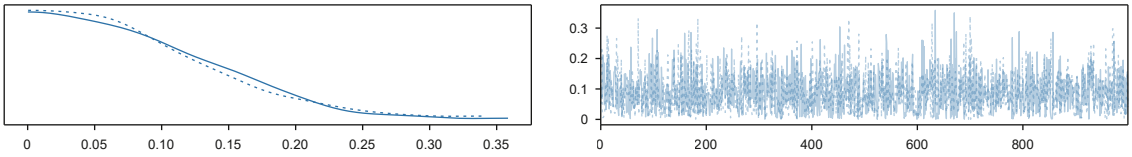


Fig. A7. Trace plots—width model (σ_{global}).

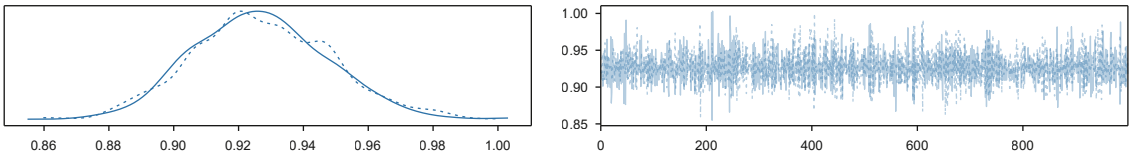


Fig. A8. Trace plots—width model (σ).

Appendix B. Autocorrelation plot for key parameters

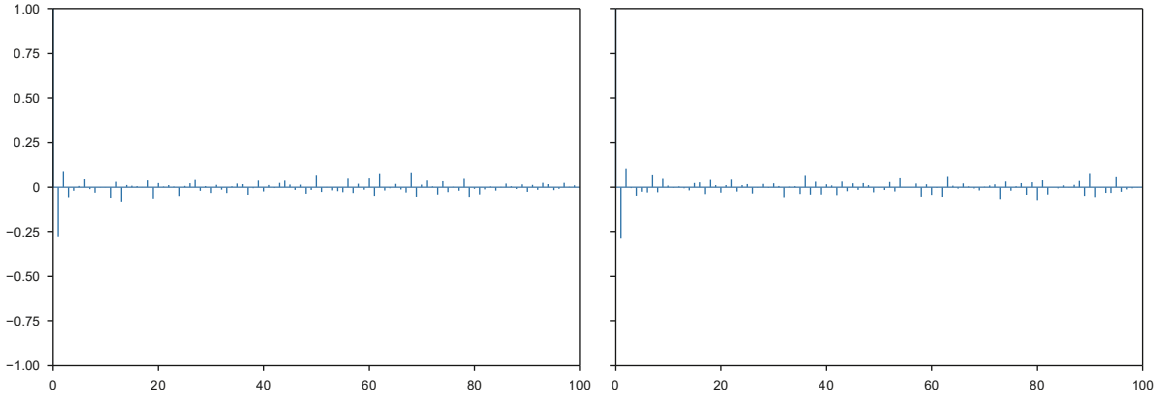


Fig. B1. Autocorrelation plot - length model (μ_{global}, two chains).

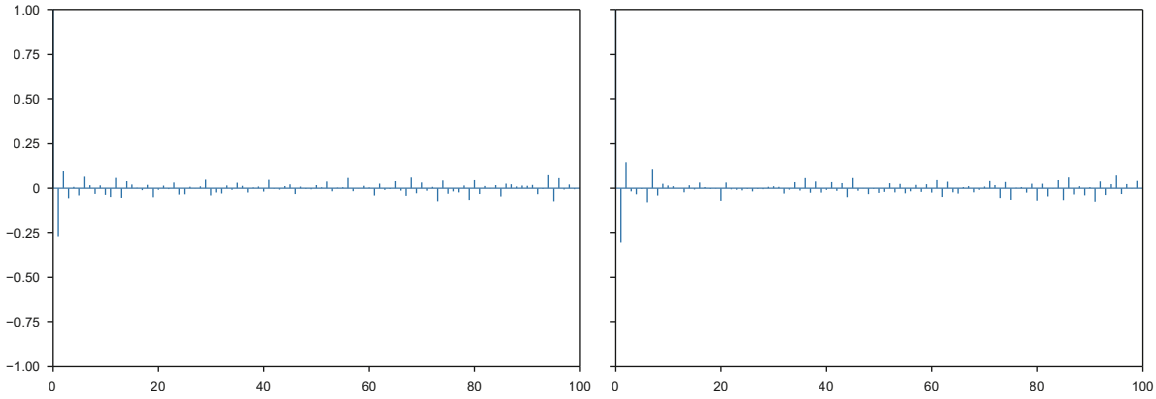


Fig. B2. Autocorrelation plots—length model (β, two chains).

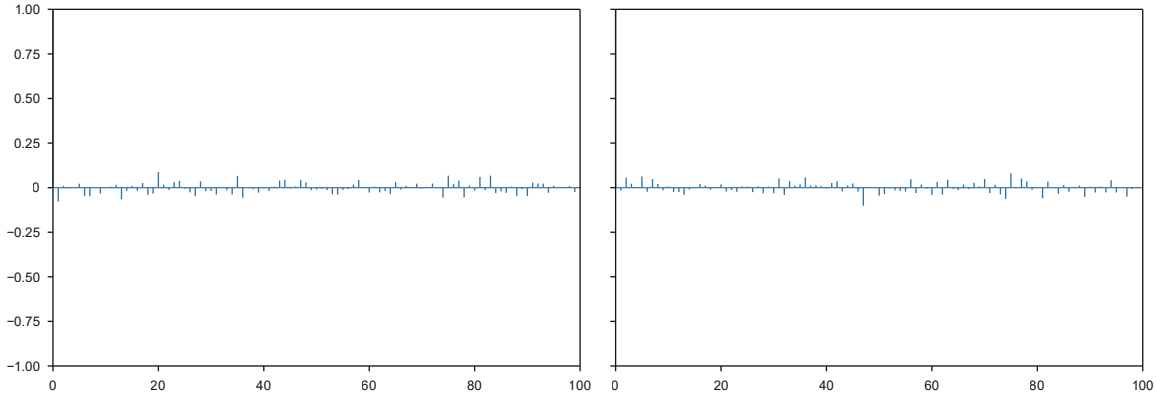


Fig. B3. Autocorrelation plots—length model (σ_{global}, two chains).

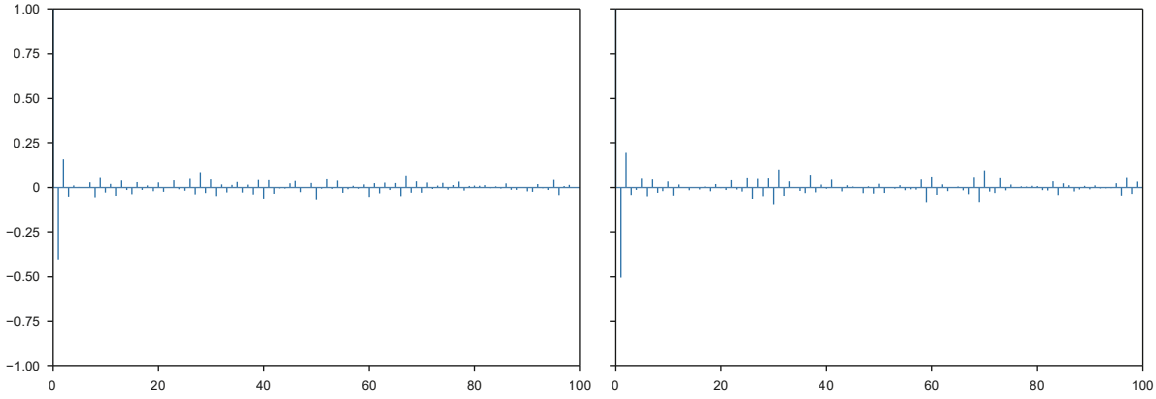


Fig. B4. Autocorrelation plots—length model (σ, two chains).

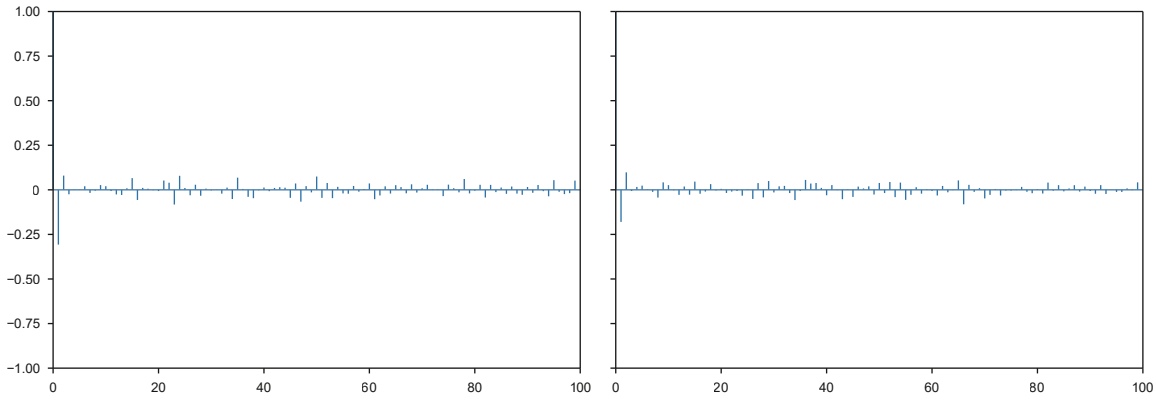


Fig. B5. Autocorrelation plots—width model (μ_{global}, two chains).

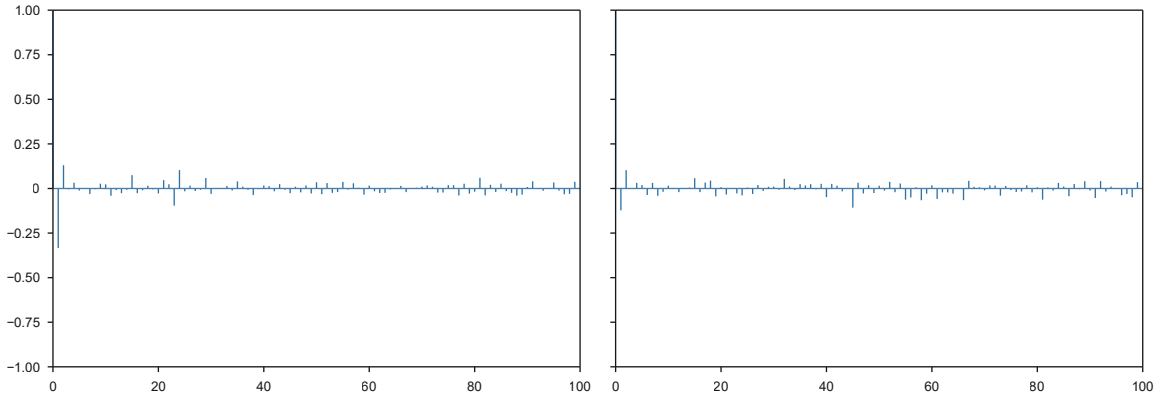


Fig. B6. Autocorrelation plots—width model (β, two chains).

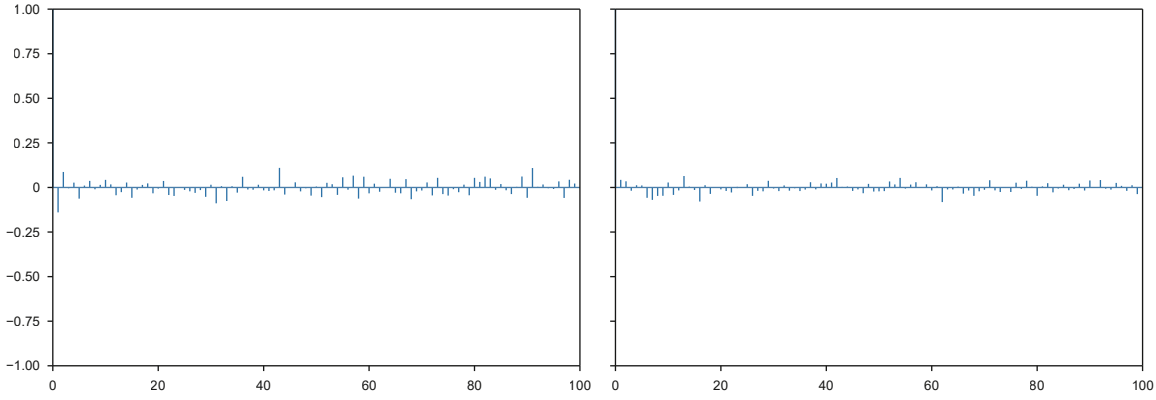


Fig. B7. Autocorrelation plots—width model (σ_{global}, two chains).

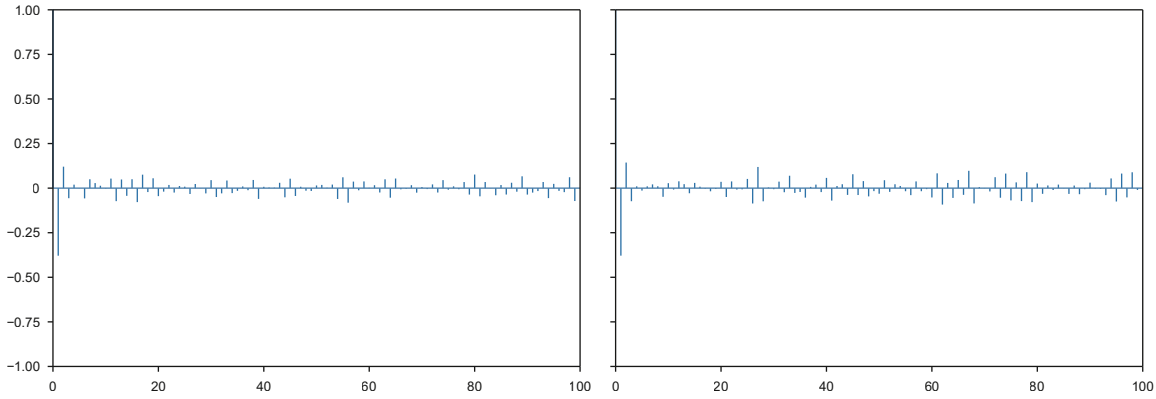


Fig. B8. Autocorrelation plots—width model (σ, two chains).

References

Ali, K., Qin, J., Faber, M.H., 2022. On information modeling in structural integrity management. Struct. Health Monit. 21 (1), 59–71. https://doi.org/10.1177/1475921720968292.

Chen, Y.F., Dong, S.H., Zang, Z.P., et al., 2021. Collapse failure and capacity of subsea pipelines with complex corrosion defects. Eng. Fail. Anal. 123, 105266. https://doi.org/10.1016/j.engfailanal.2021.105266.

Cosham, A., Hopkins, P., Macdonald, K., 2007. Best practice for the assessment of defects in pipelines–corrosion. Eng. Fail. Anal. 14 (7), 1245–1265. https://doi.org/10.1016/j.engfailanal.2006.11.035.

Cui, B.Y., Wang, H., 2024. Pipeline corrosion prediction and uncertainty analysis with an ensemble Bayesian neural network approach. Process Saf. Environ. Prot. 187, 483–494. https://doi.org/10.1016/j.psep.2024.05.011.

Der Kiureghian, A., Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter? Struct. Saf. 31 (2), 105–112. https://doi.org/10.1016/j.strusafe.2008.06.020.

El-Kady, A.H., Amin, M.T., Khan, F., et al., 2024. Analysis of CO_2 pipeline regulations from a safety perspective for offshore carbon capture, utilization, and storage (CCUS). J. Clean. Prod. 439, 140734. https://doi.org/10.1016/j.jclepro.2024.140734.

Feng, Z.Y., Su, X., Zhang, Q., et al., 2025. Probabilistic time-dependent corrosion wastage model for offshore platforms using stepwise Bayesian inference. Ocean Eng. 327, 121014. https://doi.org/10.1016/j.oceaneng.2025.121014.

Forum, P.O., 2009. Specifications and Requirements for Intelligent Pig Inspection of Pipelines. https://1library.net/document/y6mnrwgq-specifications-requirements-intelligent-pig-inspection-pipelines.html#fulltext-content.

Geyer, C.J., 1992. Practical markov chain monte carlo. Stat. Sci. 473–483. https://doi.org/10.1214/ss/1177011137.

Hu, J.L., 2024. Prediction of the internal corrosion rate for oil and gas pipelines and influence factor analysis with interpretable ensemble learning. Int. J. Pres. Ves. Pip. 212, 105329. https://doi.org/10.1016/j.ijpvp.2024.105329.

Huang, S., Wang, L., Li, Z.Y., et al., 2024. Machine learning-based prediction model for CO_2-induced corrosion on oil well cement under high-pressure and high-temperature condition. Constr. Build. Mater. 414, 134999. https://doi.org/10.1016/j.conbuildmat.2024.134999.

Hyndman, R.J., 1996. Computing and graphing highest density regions. Am. Statistician 50 (2), 120–126. https://doi.org/10.1080/00031305.1996.10474359.

Li, X.H., Zhang, L.Y., Khan, F., et al., 2021. A data-driven corrosion prediction model to support digitization of subsea operations. Process Saf. Environ. Prot. 153, 413–421. https://doi.org/10.1016/j.psep.2021.07.031.

Li, Y.T., He, X.N., Shuai, J., 2022. Risk analysis and maintenance decision making of natural gas pipelines with external corrosion based on Bayesian network. Pet. Sci. 19 (3), 1250–1261. https://doi.org/10.1016/j.petsci.2021.09.016.

Ma, H.N., Zhang, W.D., Wang, Y., et al., 2023. Advances in corrosion growth modeling for oil and gas pipelines: A review. Process Saf. Environ. Prot. 171, 71–86. https://doi.org/10.1016/j.psep.2022.12.054.

Martens, C., Botte, W., Caspeele, R., et al., 2024. Bayesian updating of relationships between crack width and corrosion level in reinforced concrete based on a large set of experimental data. Struct. Concr. 26 (5), 5743–5760. https://doi.org/10.1002/suco.202400581.

Nesic, S., Postlethwaite, J., Olsen, S., 1996. An electrochemical model for prediction of corrosion of mild steel in aqueous carbon dioxide solutions. Corrosion 52 (4), 280–294. https://doi.org/10.5006/1.3293640.

Nesic, S., Nyborg, R., Stangeland, A., et al., 2001. Mechanistic Modeling for CO_2 Corrosion with Protective Iron Carbonate Films, Nace Corrosion. NACE-01040. https://doi.org/10.5006/C2001-01040.

Nesic, S., 2007. Key issues related to modelling of internal corrosion of oil and gas pipelines – A review. Corros. Sci. 49 (12), 4308–4338. https://doi.org/10.1016/j.corsci.2007.06.006.

Nesic, S., 2012. Effects of multiphase flow on internal CO_2 corrosion of mild steel pipelines. Energy Fuel. 26 (7), 4098–4111. https://doi.org/10.1021/ef3002795.

Papavinasam, S., Revie, R.W., I.Friesen, W., et al., 2006. Review of models to predict internal pitting corrosion of oil and gas pipelines. Corrosion Rev. 24 (3–4), 173–230. https://doi.org/10.1515/CORRREV.2006.24.3-4.173.

Qin, J.J., 2022. Preposterior analysis considering uncertainties and dependencies of information relevant to structural performance. ASCE-ASME J. Risk Uncert. Eng. Syst. Part A: Civ. Eng. 8 (1), 04021085. https://doi.org/10.1061/AJRUA6.0001211.

Qin, M., Chen, S.J., Ye, N., et al., 2025. Study on the evolution mechanism and influencing factors of CO_2 corrosion product film under pipe flow. Int. J. Pres. Ves. Pip. 216, 105506. https://doi.org/10.1016/j.ijpvp.2025.105506.

Seghier, M.E.A.B., Keshtegar, B., Tee, K.F., et al., 2020. Prediction of maximum pitting corrosion depth in oil and gas pipelines. Eng. Fail. Anal. 112, 104505. https://doi.org/10.1016/j.engfailanal.2020.104505.

Seghier, M.E.A.B., Höche, D., Zheludkevich, M., 2022. Prediction of the internal corrosion rate for oil and gas pipeline: implementation of ensemble learning techniques. J. Nat. Gas Sci. Eng. 99, 104425. https://doi.org/10.1016/j.jngse.2022.104425.

Simonsen, K.R., Goebel, J., Hansen, D.S., et al., 2025. The influence of temperature, H_2O, and NO_2 on corrosion in CO_2 transportation pipelines. Process Saf. Environ. Prot., 107190 https://doi.org/10.1016/j.psep.2025.107190.

Song, D.D., Zhou, Q., Xu, D., et al., 2025. Corrosion prediction and factors analysis of 2A12 aluminum alloy in marine environment based on data mining. Mater. Today Commun. 42, 111324. https://doi.org/10.1016/j.mtcomm.2024.111324.

Sun, M., Wang, X.H., Cui, W., et al., 2025. Effect of temperature on corrosion of L245 steel under CO_2-SRB corrosion system. Microorganisms 13 (3), 500. https://doi.org/10.3390/microorganisms13030500.

Tian, H.F., Guan, F., Wan, F., et al., 2024. The effect of corrosion defects on the collapse pressure of submarine pipelines. Ocean Eng. 310, 118647. https://doi.org/10.1016/j.oceaneng.2024.118647.

Wang, Y.H., Xia, A.L., Zhang, P., et al., 2022. Probabilistic physical modeling of randomly corroded surface and its use in reliability analysis of corroded pipelines under spatiotemporal vibration. Ocean Eng. 262, 112219. https://doi.org/10.1016/j.oceaneng.2022.112219.

Yuan, Y., Li, C., Zhao, Y.Y., et al., 2025. Crevice corrosion mechanism of L80-13Cr in Cl-containing supercritical CO_2 water-rich phase considering the influence of SO_2. J. Supercrit. Fluids 222, 106577. https://doi.org/10.1016/j.supflu.2025.106577.

Yue, Y.G., Yin, Z.B., Li, S.M., et al., 2025. Corrosion behavior of mild steel in various environments including CO_2, H_2S, and their combinations. Metals 15 (4), 440. https://doi.org/10.3390/met15040440.