

Article

A Multi-Graph Attention Fusion Network for Dam Deformation Prediction Under Data Missing Conditions

Weiting Lu ^{1,2}, Dongjie Wu ^{3,*}, Jian Liang ^{1,2}, Guanghe Zhang ^{1,2}, Zhenhao Wu ⁴ and Na Xia ⁴¹ Anhui & Huaihe River Institute of Hydraulic Research, Hefei 230088, China² Anhui Provincial Key Laboratory of Water Science and Intelligent Water Conservancy, Hefei 230088, China³ College of Automotive and Transportation Engineering, Hefei University of Technology, Hefei 230009, China⁴ College of Computer and Information Science, Hefei University of Technology, Hefei 230601, China

* Correspondence: dongjie@mail.hfut.edu.cn; Tel.: +86-178-8015-2053

Abstract

Dam deformation monitoring is essential for ensuring the safe operation of hydraulic structures, yet practical data are often compromised by missing values and noise, while spatial coupling among monitoring points further complicates prediction. To address these challenges, this study proposes a Spatio-Temporal Multi-Graph Attention Fusion Network (STMGA FN) for dam deformation prediction and risk early warning under incomplete data conditions. Data quality is enhanced through a data-quality-aware hierarchical adaptive imputation mechanism combined with a VMD–wavelet joint denoising strategy. A multi-graph spatial modeling framework integrating temporal similarity, spatial proximity, structural zoning, and measuring-line connectivity is constructed, and fuses multi-source spatial features through a lightweight adaptive attention mechanism. A parameter-sharing recursive probabilistic temporal modeling approach is adopted to jointly predict deformation values and their associated uncertainties. Based on the predicted confidence intervals, a four-level risk classification and early-warning scheme is established. Experimental results on real GNSS monitoring data from dam sites demonstrate that the proposed method achieves an *RMSE* of 0.3588 mm, an *MAE* of 0.1738 mm, and an R^2 of 0.9865, outperforming baseline models including LSTM, TCN, CNN-LSTM, and STGCN. Moreover, the correlation between predictive uncertainty and actual error reaches 0.892, verifying the effectiveness and reliability of the proposed method for dam safety monitoring under complex conditions.

Keywords: dam deformation prediction; multi-source data fusion; Spatio-Temporal Multi-Graph modeling; probabilistic time-series forecasting; risk early warning



Academic Editor: Ahmed Abu-Siada

Received: 1 February 2026

Revised: 26 March 2026

Accepted: 27 March 2026

Published: 31 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Dams are critical infrastructure for water resource management, flood control, and hydroelectric power generation. They play a key role in protecting downstream communities and supporting socio-economic stability. As dams age and environmental loads change, structural deformation gradually builds up. Abnormal deformation is often seen as an early sign of structural instability or failure. Therefore, accurate and reliable deformation prediction is important for ensuring dam safety [1]. In recent years, GNSS, total stations, and automated monitoring systems have been widely used. These systems allow continuous collection of multi-point, long-term monitoring data, which supports data-driven safety assessment [2]. At the same time, deep learning methods for time-series modeling have

developed rapidly, offering new ways to capture complex nonlinear patterns [3]. However, stable and reliable prediction under complex noise, non-stationarity, and multi-source data coupling remains challenging [4].

Traditional dam deformation prediction methods mainly include the Hydrostatic–Seasonal–Time (HST) model, multiple regression models, and ARIMA time-series models. These methods predict deformation by building explicit relationships with factors such as water level, temperature, and aging. They offer some physical interpretability. However, they have difficulty capturing complex nonlinear features and spatial coupling among multiple monitoring points [5]. With the development of deep learning, models such as LSTM, GRU, TCN, and attention-based networks have been applied to dam deformation prediction. These models have shown good results in single-point or few-point settings [6]. However, they mainly focus on temporal modeling and often treat each monitoring point as an independent sequence. Only a few studies have started to consider spatio-temporal data fusion [7]. As a result, the spatial correlations among multiple monitoring points are not fully used.

Graph Neural Networks (GNNs) are effective tools for modeling spatial dependencies in multi-node systems. In recent years, spatio-temporal methods that combine graph structures with temporal networks have been applied to structural health monitoring. These methods show good potential for multi-point modeling [8]. However, most existing studies rely on a single predefined graph structure. A single graph cannot fully represent multi-source spatial relationships, such as geometric proximity, deformation similarity, and structural zoning. There is still room to improve spatial representation. Graph modeling is also moving from predefined rule-based graphs toward data-driven adaptive graph learning. In addition, real monitoring data often have missing values, noise, and non-stationarity. Deep learning models are sensitive to data quality. Directly modeling low-quality data often leads to poor prediction performance and stability. Furthermore, most existing models only produce deterministic point predictions. They cannot quantify prediction uncertainty, which limits their use in safety assessment and risk decision-making [9].

Based on existing research, intelligent dam deformation prediction still faces four key issues. First, the spatial relationships among monitoring points are multi-source and heterogeneous. A single graph structure cannot fully capture these complex spatial features. Second, missing values and noise in monitoring data are unavoidable, and existing methods are not robust when data quality drops. Third, temporal prediction models tend to accumulate errors in long-term rolling predictions. The gap between training and inference phases makes this problem worse. Fourth, most models do not quantify prediction uncertainty, which limits their ability to support risk warnings in safety-critical scenarios. These issues show the need for a spatio-temporal prediction framework that can model multi-source spatial relationships, handle data quality problems, maintain long-term prediction stability, and support uncertainty quantification.

To address these challenges, this paper proposes a spatio-temporal prediction framework for dam multi-point deformation. The framework combines multi-source spatial relationships, temporal features, and environmental factors for joint modeling. In the spatial dimension, it uses multiple complementary graph structures to capture coupling among monitoring points. In the temporal dimension, it models the dynamic evolution of deformation. Environmental factors such as water level, temperature, and rainfall are also included to reflect the effects of external conditions on structural response. To handle missing values and noise in monitoring data, a data-quality-aware hierarchical adaptive imputation and denoising strategy is introduced. This provides stable data input for the model. To improve long-term prediction stability, the framework uses a parameter-sharing temporal decoding structure. This structure reduces error accumulation in rolling predictions. In addition, a

probabilistic prediction mechanism is introduced to quantify uncertainty, providing a basis for risk assessment.

Focusing on spatial correlation modeling, data quality robustness, and safety decision-making requirements for dam deformation prediction in real engineering scenarios, the main contributions of this paper are as follows:

- (1) An integrated modeling framework combining spatio-temporal prediction and risk warning with uncertainty quantification is constructed, achieving closed-loop support from deformation prediction to risk decision-making through probabilistic prediction and adaptive warning mechanisms;
- (2) A spatio-temporal collaborative modeling method integrating multi-source spatial topological relationships, temporal dynamic features, and environmental driving factors is proposed, realizing joint prediction under multi-point and multi-variable conditions, with a lightweight adaptive multi-graph fusion mechanism designed to improve modeling efficiency for complex spatial relationships;
- (3) A data-quality-aware hierarchical adaptive imputation and denoising strategy is introduced, in which imputation methods are dynamically selected according to the temporal characteristics of missing data to effectively mitigate the interference caused by missing values and noise, thereby providing a stable and reliable data input foundation for spatio-temporal modeling. Systematic experiments on real dam deformation monitoring datasets validate the advantages of the proposed method in prediction accuracy, robustness, and risk warning effectiveness.

In summary, this study proposes a spatio-temporal collaborative prediction framework based on multi-source heterogeneous information for dam deformation forecasting. The proposed framework jointly captures the complex spatial coupling relationships among multiple monitoring points, the temporal evolution patterns of deformation, and the effects of environmental driving factors. By incorporating data-quality-aware processing and probabilistic prediction, it further enhances the stability and practical applicability of the model under real-world monitoring conditions. Overall, the study provides an effective and intelligent solution for improving prediction accuracy and robustness, while also offering stronger support for safety monitoring, risk identification, and engineering decision-making.

The remainder of this paper is organized as follows. Section 2 reviews the related work. Section 3 presents the data-quality-aware hierarchical adaptive processing strategy and the proposed spatio-temporal multi-graph attention fusion network. Section 4 reports the experimental results and comparative analysis. Section 5 discusses the engineering significance, limitations, and future directions of the proposed method. Finally, Section 6 concludes the paper.

2. Related Work

In engineering practice, dam deformation prediction has long relied on mechanism–statistical hybrid models. These models predict deformation by describing explicit relationships with variables such as water level, temperature, and aging. Recent studies have continued to improve this approach. For example, time-lag effects have been introduced to better reflect the delayed influence of temperature on deformation. Information-theoretic indicators have also been used to describe relationships among variables, improving model adaptability and interpretability under real conditions [1,10]. These methods have clear structure and strong engineering interpretability. However, they often fail to achieve stable high-precision predictions in complex scenarios with strong nonlinearity, multi-source driving, and multi-point coupling.

With the growth in monitoring data scale and dimensionality, deep learning has become an important approach for dam deformation prediction. In recent years, many

studies have used hybrid frameworks of “signal decomposition/denoising + deep network” to reduce non-stationarity and noise in monitoring sequences. For example, CEEMDAN has been combined with self-attention LSTM to better capture periodic and nonlinear patterns [11]. Other studies use VMD-based denoising to improve model robustness and prediction accuracy through parameter optimization and attention structures [12]. Some work proposes new modal decomposition strategies combined with deep learning to further improve modeling of complex deformation signals [13]. In dam deformation prediction, combined models such as VMD/CNN/LSTM and CNN/GRU have been used for feature extraction and long-term sequence modeling. These studies confirm the effectiveness of deep learning for complex deformation time series [14,15]. Overall, deep learning shows outstanding performance in single-point or few-point predictions. However, it still faces limitations in multi-point collaborative modeling, cross-point spatial coupling, and stability under missing data conditions.

To capture spatial dependencies among multiple monitoring points, spatio-temporal modeling has evolved from independent multi-sequence modeling to joint spatial-temporal modeling. In recent years, studies on dam displacement prediction have used graph convolution or attention mechanisms to jointly learn spatial relationships and temporal dynamics among monitoring points. This has improved long-term prediction performance and spatial consistency [16]. In 2025, a framework based on VAE and Spatio-Temporal Graph Convolutional Networks (STGCN) was proposed for multi-point displacement monitoring. It better captures spatial relationships among monitoring points and improves prediction performance [17]. However, in dam monitoring, most existing studies still rely on a single predefined graph (e.g., distance-based or correlation-based) or a single learning mechanism. It is difficult for a single graph to represent multi-source spatial relationships such as geometric proximity, deformation similarity, and structural zoning at the same time.

In the broader fields of spatio-temporal prediction and engineering monitoring, graph construction has evolved from predefined rule-based graphs toward data-driven adaptive graph learning. Related studies have shown that dynamic graph convolution, spatio-temporal attention, and learnable adjacency relationships can help discover hidden spatial dependencies from multi-sensor data. This improves the representation of complex structural states. For example, one study [18] combined dynamic graph convolutional networks with a Transformer for structural state identification. Another study [19] proposed the U-GraphFormer framework, which uses a dynamic spatio-temporal graph learning mechanism for structural damage detection and severity assessment. For dam monitoring, which involves multiple prior relationships such as geometric proximity, structural zoning, and response similarity, multi-graph fusion and adaptive relationship modeling provide a reasonable direction for multi-point deformation prediction.

In real monitoring systems, equipment failures, communication interruptions, and external interference often cause missing values and outliers. These problems directly affect prediction model reliability. Research on missing data in dam monitoring has increased in recent years. On one hand, multi-head attention and Generative Adversarial Networks (GAN) have been applied to dam sensor data imputation to improve recovery quality under complex missing patterns [20]. However, these deep generative and attention-based models usually require complex network architectures and iterative training. This leads to high computational costs. They also have difficulty incorporating engineering prior knowledge. On the other hand, some studies have proposed frameworks that handle missing values and prediction targets together from an “imputation + behavior prediction/label discrimination” perspective to improve engineering applicability [21]. In the related field of surface deformation monitoring, “spatio-temporal graph attention” frameworks have been designed for gap filling, with a focus on maintaining physical deformation patterns in

the imputed results [22]. Overall, missing data processing has evolved from simple interpolation toward structure preservation, end-to-end learning, and combined spatio-temporal correlation. However, systematic solutions that integrate missing data handling, denoising, and prediction modeling for dam multi-point deformation prediction are still lacking.

Dam safety is a typical safety-critical application. Providing only point predictions cannot meet risk decision-making needs. In recent years, the Structural Health Monitoring (SHM) field has placed more emphasis on expressing uncertainty in probabilistic terms. Bayesian frameworks and probabilistic inference have been used to better capture environmental variation and model error [23]. At the same time, data-driven modeling of the “probabilistic structure of prediction errors” has been used to analyze how errors affect downstream tasks such as damage detection [24]. In more general deep learning time series research, recent surveys point out that long-term prediction, distribution drift, and reliability of uncertainty estimation are still key challenges for deploying probabilistic prediction in engineering [25]. Therefore, this study combines uncertainty quantification with multi-point spatio-temporal modeling and data quality assurance. It also connects with hierarchical risk warning to support deformation prediction and safety decision-making under complex conditions.

3. Methodology

3.1. Problem Definition

Dam deformation prediction is defined as a spatio-temporal prediction problem with uncertainty quantification for risk decision-making. Given historical deformation measurements from N monitoring points over T time steps $X \in \mathbb{R}^{T \times N \times D}$ and environmental covariates $V \in \mathbb{R}^{T \times M}$ (including water level, temperature, and rainfall, $M = 3$), the goal is to predict deformation $\hat{Y} \in \mathbb{R}^{T' \times N}$ and its uncertainty $\sigma \in \mathbb{R}^{T' \times N}$ for the next T' time steps.

This study learns a probabilistic mapping [26] that predicts both the mean deformation μ and the standard deviation σ for future time steps at each monitoring point, based on historical observations and environmental covariates. The model provides both point forecasts and uncertainty estimates, supporting risk-informed decision-making in dam safety management. The overall research process of this study is shown in Figure 1.



Figure 1. Overall research workflow of the Spatio-Temporal Multi-Graph Adaptive Fusion Network (STMGAFN)-based framework: from raw GNSS monitoring data to probabilistic prediction and four-level risk warning.

3.2. Data-Quality-Aware Hierarchical Adaptive Processing

To handle the multi-scale temporal characteristics of missing data in dam monitoring time series, a data-quality-aware hierarchical adaptive processing strategy is proposed. The strategy automatically identifies the data quality state based on the duration of missing intervals. It then switches among three imputation methods: cubic spline interpolation for short-term gaps (≤ 3 days) to preserve temporal continuity, bidirectional GBR regression for long-term gaps (> 3 days) to use contextual information, and spatial neighborhood weighted interpolation for boundary missing segments. This approach overcomes the limitations of single-strategy methods and improves both processing efficiency and effectiveness. This strategy is not just a conventional preprocessing step. It is a key component of the proposed framework for handling complex data-quality issues in monitoring data.

3.2.1. Hierarchical Imputation Strategy

A hierarchical imputation strategy is proposed by combining temporal interpolation, machine learning regression, and spatial interpolation methods according to the duration of data gaps, thereby balancing physical plausibility with predictive accuracy.

For short-term gaps (≤ 3 days), cubic spline interpolation is used. It provides smooth and physically reasonable estimates by constructing piecewise cubic polynomials between adjacent observations:

$$S(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, x \in [x_i, x_{i+1}] \quad (1)$$

The coefficients $\{a_i, b_i, c_i, d_i\}$ are determined by satisfying continuity conditions for function values as well as first- and second-order derivatives:

$$S(x_i) = y_i, S'(x_i^-) = S'(x_i^+), S''(x_i^-) = S''(x_i^+) \quad (2)$$

This method exhibits high computational efficiency, effectively suppresses high-frequency noise, and maintains the smooth evolutionary characteristics of deformation time series, consistent with the actual physical behavior of dam structures.

For long-term missing segments (> 3 days), interpolation alone cannot capture the trend of the time series. Therefore, a bidirectional recursive GBR imputation model combined with distance weighting is introduced. Note that bidirectional GBR imputation is applied only to missing values within the training period and does not involve any information from the validation or test sets. For missing values in the validation and test periods, only a causal forward imputation strategy based on historical observations is adopted, thereby preserving temporal causality during evaluation.

Forward and backward prediction models are constructed separately:

$$f_{forward}(t) = GBR(X_{t-k:t-1}, V_{t-k:t-1}) \quad (3)$$

$$f_{backward}(t) = GBR(X_{t+1:t+k}, V_{t+1:t+k}) \quad (4)$$

where X denotes the monitored deformation variables and V represents environmental covariates. The final imputed value is obtained through adaptive weighted fusion:

$$\hat{X}_t = \alpha f_{forward}(t) + (1 - \alpha) f_{backward}(t) \quad (5)$$

where the weighting coefficient α is defined according to the temporal distance between the current time step and the nearest available observation.

$$\alpha = \frac{\exp(-d_{forward})}{\exp(-d_{forward}) + \exp(-d_{backward})} \quad (6)$$

where $d_{forward}$ and $d_{backward}$ are the temporal distances to the nearest valid observations in the forward and backward directions. This weighting mechanism gives more weight to observations that are closer in time. This improves the stability and reliability of long-term missing data imputation. In this study, the short-missing threshold is set to 3 days based on two considerations. First, under short-term missing conditions, deformation time series usually retain strong temporal continuity and are therefore suitable for smooth completion using cubic spline interpolation. Second, in the present dataset, missing segments are mainly concentrated in isolated 1-day gaps and continuous gaps longer than 7 days; thus,

the 3-day threshold can naturally distinguish short-term occasional missingness from relatively long continuous missingness.

When temporal context is not available, such as at data boundaries, spatial interpolation based on neighboring monitoring points is used as a backup strategy. Due to the spatial heterogeneity within the dam structure, spatial imputation does not use all monitoring points. Instead, only valid observations from the k -nearest neighbors of the target point are used for local weighted estimation:

$$\hat{X}_i = \frac{\sum_{j \in \mathcal{N}_k(i)} \omega_{ij} X_j}{\sum_{j \in \mathcal{N}_k(i)} \omega_{ij}} \quad \omega_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \quad (7)$$

where $\mathcal{N}_k(i)$ denotes the set of k nearest neighbors of monitoring point i with $k = 3$. The weight ω_{ij} is determined by the spatial distance d_{ij} between monitoring points and is computed using a Gaussian kernel function. This strategy preserves the physical meaning of distance information and avoids the influence of structurally distant points. This improves the robustness and reliability of interpolation. Spatial interpolation serves as a fallback strategy when temporal interpolation is not applicable. Its purpose is to provide locally stable estimates in boundary scenarios.

3.2.2. VMD–Wavelet Joint Denoising

GNSS monitoring signals are typically affected by three types of interference: low-frequency instrument drift, environmentally driven non-stationary responses, and high-frequency random noise. To further improve data quality, this study proposes a two-stage denoising method that combines Variational Mode Decomposition (VMD) with wavelet thresholding. This method balances signal decomposition adaptivity with noise suppression.

Stage 1: Variational Mode Decomposition (VMD)

First, the VMD method is applied to decompose the raw dam displacement signal into K Intrinsic Mode Functions (IMFs) with finite bandwidth by solving the following variational optimization problem:

$$\min_{\{\mu_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \quad s.t. \sum_{k=1}^K \mu_k(t) = x(t) \quad (8)$$

where $\mu_k(t)$ represents the k -th modal component with center frequency ω_k . In the present experiments, the number of modes was set to $K = 3$. This choice was not determined solely by minimizing reconstruction errors but was selected based on a sensitivity analysis over $K = 2 \rightarrow 8$, with joint consideration of denoising strength, signal fidelity, modal energy distribution, and physical interpretability. (see Appendix A).

VMD can effectively separate components at different frequency scales. It is particularly suitable for identifying drift and slow-varying trends. However, it does not fully suppress the residual random noise within individual IMF components. Further denoising is still needed.

Stage 2: Wavelet Threshold Denoising

The IMF components are first subjected to wavelet transform to obtain corresponding wavelet coefficients ρ which are then processed through soft thresholding shrinkage:

$$\eta_{soft}(\rho, \lambda) = \text{sgn}(\rho) \cdot \max(|\rho| - \lambda, 0) \quad (9)$$

where the threshold λ is determined by the universal threshold rule:

$$\lambda = \sigma_n \sqrt{2 \log N} \quad (10)$$

where N denotes the signal length and σ_n represents the noise standard deviation. It should be noted that, in this study, wavelet threshold denoising is applied to the high-frequency modes obtained after VMD. Since the high-frequency noise components are concentrated after VMD-based pre-separation, the finest-scale wavelet coefficients are empirically closer to an approximately Gaussian distribution (see Appendix B). Therefore, the universal threshold is adopted for threshold determination, and MAD is used to obtain a robust estimate of the noise scale.

Subsequently, inverse wavelet transform and reconstruction yield the final denoised signal:

$$\tilde{x}(t) = \sum_{k=1}^K WT^{-1}[\eta(WT[\mu_k(t)])] \quad (11)$$

where $WT(\cdot)$ and $WT^{-1}(\cdot)$ denote the wavelet transform and its inverse, respectively.

Finally, to enhance the numerical stability and convergence efficiency of subsequent model training, Min-Max normalization is applied to scale the data to the $[0, 1]$ interval:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (12)$$

where x denotes the original deformation feature, and x_{min} and x_{max} represent the minimum and maximum values of that feature in the training set, respectively. The normalization parameters $\{x_{min}, x_{max}\}$ are computed solely based on the training set and preserved for inverse transformation during inference.

Wavelet thresholding is effective in attenuating high-frequency random noise, but when directly applied to the original mixed GNSS signal alone, it is difficult to simultaneously account for the differences among low-frequency drift, non-stationary responses, and high-frequency noise in terms of temporal scales and statistical characteristics.

Therefore, a two-stage VMD–Wavelet strategy is used. First, VMD decomposes the complex non-stationary signal into modes. This achieves an initial separation of components at different scales. Then, wavelet threshold denoising is applied to each IMF component to further suppress the remaining random noise. This strategy combines the strength of VMD in non-stationary signal decomposition with the effectiveness of wavelet thresholding in noise suppression. It handles both instrument drift and random noise while preserving the true deformation-response characteristics. The frequency-spectrum comparison before and after denoising shows that this method suppresses high-frequency noise without significantly changing the main characteristics of the signal.

3.3. Multi-Source Spatial Topology Modeling

The spatial relationships among dam monitoring points involve multiple factors. A single graph structure is not enough to represent them. To address this, multi-source heterogeneous spatial graphs are constructed from four perspectives: temporal similarity, spatial proximity, structural partitioning, and measurement-line connectivity. These graphs together provide a more complete representation of spatial associations.

(1) Temporal Similarity-Based Adjacency Graph

To characterize the similarity in deformation patterns among different monitoring points, Dynamic Time Warping (DTW) is introduced to measure the similarity between

time series. Given time series s_i and s_j for monitoring points i and j , the DTW distance is defined as:

$$DTW(s_i, s_j) = \min_{\pi} \sum_{(p,q) \in \pi} d(s_i[p], s_j[q]) \quad (13)$$

where π represents an alignment path satisfying continuity and boundary constraints. Based on DTW distance, the adjacency matrix is constructed as:

$$A_{ij}^{(1)} = \begin{cases} \frac{1}{DTW(s_i, s_j)}, & \text{if } \frac{1}{DTW(s_i, s_j)} \geq \theta_{DTW} \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

The threshold θ_{DTW} is determined in a data-driven manner, specifically as the 25th percentile of all nonzero DTW similarity values. This choice aims to balance graph connectivity and the filtering of weak similarity edges (see Appendix C). In addition, to reduce computational cost, an approximate DTW algorithm with a Sakoe–Chiba band constraint is adopted, which reduces the complexity for a single sequence pair from $O(T^2)$ to $O(T \times W)$, where the window width is set to $W = 30$.

(2) Spatial Proximity-Based Adjacency Graph

Considering that neighboring monitoring points in a continuous stress field may experience similar deformation responses, spatial adjacency relationships are constructed based on the three-dimensional spatial coordinates of monitoring points. Let d_{ij} denote the Euclidean distance between monitoring points i and j :

$$d_{ij} = \|loc_i - loc_j\|_2 \quad (15)$$

The adjacency matrix is defined using a Gaussian kernel function:

$$A_{ij}^{(2)} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma_d^2}\right), & \text{if } \exp\left(-\frac{d_{ij}^2}{\sigma_d^2}\right) \geq \theta_{Gauss} \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

where σ_d denotes the average Euclidean distance between all monitoring points, i.e., $\sigma_d = \text{mean}(d_{ij})$. Here, d_{ij} denotes the Euclidean distance between monitoring points i and j .

(3) Structural Partition-Based Consistency Adjacency Graph

Considering the differences in loading and constraint conditions across different structural partitions of the dam, if monitoring points i and j belong to the same structural partition, a direct connection is assumed to exist between them, with the adjacency relationship defined as:

$$A_{ij}^{(3)} = \begin{cases} 1, & \text{if } \text{partition}_i = \text{partition}_j \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

(4) Measurement Line Connectivity-Based Adjacency Graph

Monitoring points are typically deployed along different elevation measurement lines, with points on the same line experiencing similar loading conditions. The adjacency matrix is constructed based on measurement line identifiers:

$$A_{ij}^{(4)} = \begin{cases} 1, & \text{if } \text{line}_i = \text{line}_j \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Additionally, to ensure that the GCN can fully leverage each node's own feature information, self-loops are added to all adjacency matrices prior to the graph convolution, i.e., $\tilde{A} = A + I$, where I is the identity matrix. Symmetric normalization is then applied:

$$\tilde{A}^{(k)} = D^{-1/2} A^{(k)} D^{-1/2}, \quad k = 1, \dots, 4 \quad (19)$$

where D represents the degree matrix of the corresponding adjacency matrix A .

In this study, the 23 GNSS monitoring stations are divided into six structural partitions and four survey lines based on their spatial locations and engineering structural characteristics.

Structural partitions: The F-series stations are grouped according to their relative positions along the dam into upstream (F3, F11, F13, F17, F19, F23, F25, F33; 8 stations) and downstream (F12, F4, F18, F14, F24, F20, F34, F26; 8 stations) groups. The Y-series stations are divided based on spatial location into four partitions: upper-left (Y50, Y52), lower-left (Y38, Y40), upper-right (Y49), and lower-right (Y37, Y39).

Measurement line grouping: Based on elevation characteristics, the stations are grouped into four measurement lines: Line 1 (upstream F-series, elevation ~27.5–27.7 m), Line 2 (downstream F-series, elevation ~28.1–28.3 m), Line 3 (upper Y-series, elevation ~24.6–24.7 m), and Line 4 (lower Y-series, elevation ~25.5–25.7 m).

3.4. Spatio-Temporal Multi-Graph Adaptive Fusion Network (STMGAFN)

To fully exploit multi-source spatial topological structures and characterize the temporal evolutionary patterns of dam deformation, this study proposes a Spatio-Temporal Multi-Graph Adaptive Fusion Network (STMGAFN). The overall model architecture is illustrated in Figure 2, comprising four core modules: multi-graph convolutional feature extraction, lightweight adaptive fusion, temporal modeling, and probabilistic output.

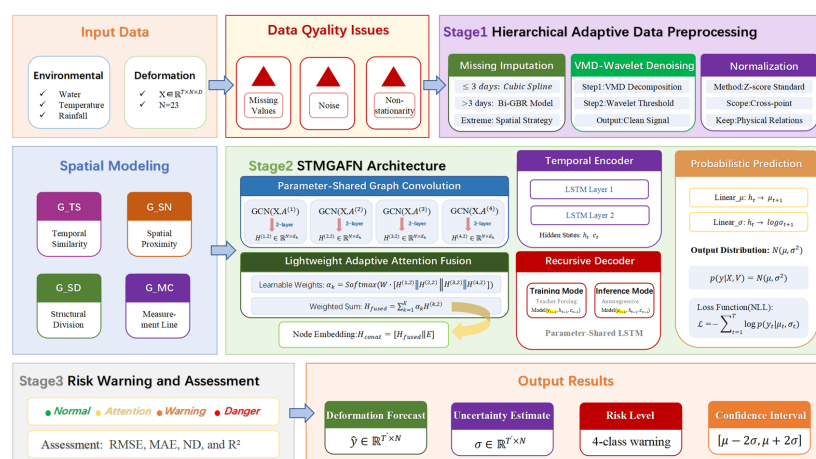


Figure 2. Overall framework of the proposed Spatio-Temporal Multi-Graph Adaptive Fusion Network (STMGAFN) for GNSS-based dam deformation prediction and risk warning. The framework starts with environmental and raw deformation data and addresses major data-quality issues, including missing values, noise, and non-stationarity. Stage 1 performs hierarchical adaptive preprocessing, including missing-value imputation, VMD-wavelet denoising, and normalization. Stage 2 constructs four spatial graphs, namely temporal similarity G_{TS} , spatial proximity G_{SN} , structural division G_{SD} , and measurement-line connectivity G_{MC} , and integrates them through the STMGAFN for spatio-temporal feature learning and probabilistic prediction. Stage 3 converts the predicted mean and uncertainty into confidence intervals and four warning levels. Colored boxes denote functional modules, arrows indicate the data-processing and information-flow paths between stages, and symbols such as μ and σ represent the predicted mean and uncertainty, respectively.

3.4.1. Multi-Graph Convolution and Lightweight Adaptive Fusion

At each time step, the input is denoted as $X_t \in \mathbb{R}^{N \times (1+M+1)}$, which consists of deformation measurements, environmental covariates (including water level, temperature, and rainfall), and monitoring point indices.

Assume that there are K graph views ($K = 4$), where each graph $G^{(k)}$ corresponds to a normalized adjacency matrix $\tilde{A}^{(k)}$. For each graph, a two-layer graph convolutional structure is employed to extract spatial representations. The propagation process of the k -th graph at the first and second layers is formulated as:

$$H^{(k,1)} = \text{ReLU}\left(\tilde{A}^{(k)} X W_1^{(k)}\right) \quad (20)$$

$$H^{(k,2)} = \text{ReLU}\left(\tilde{A}^{(k)} H^{(k,1)} W_2^{(k)}\right) \quad (21)$$

where $W_1^{(k)} \in \mathbb{R}^{(1+D_c) \times d_h}$, $W_2^{(k)} \in \mathbb{R}^{d_h \times d_h}$ are learnable parameters, d_h denotes the hidden dimension, and D_c represents the covariate dimension ($D_c = M$). Ultimately, each graph outputs a spatial representation $H^{(k,2)} \in \mathbb{R}^{N \times d_h}$.

In multi-graph information fusion problems, traditional methods usually use simple addition and concatenation, which cannot distinguish the importance of different graphs; the number of parameters usually increases substantially, and the complexity often becomes $O(Kd_h^2)$, which is not suitable for real-time applications. To address this, a method is proposed that uses a lightweight fusion mechanism that scales with the number of parameters $O(K)$, and achieves effective multi-graph fusion by assigning learnable weights to each graph.

First, a learnable weight set $w = \{w_1, w_2, \dots, w_K\}$ is introduced for the K graphs, with Softmax normalization applied to obtain fusion coefficients:

$$\alpha_k = \frac{\exp(w_k)}{\sum_{j=1}^K \exp(w_j)} \quad (22)$$

Subsequently, weighted summation of the spatial representations from each graph yields the fused result:

$$H_{fused} = \sum_{k=1}^K \alpha_k H^{(k,2)} \quad (23)$$

The learned weights α_k provide interpretability, directly indicating the relative importance of each graph structure. This module introduces only one learnable scalar per graph in the fusion stage. It should be noted that the $O(K)$ complexity refers solely to the parameters of the fusion mechanism, while the overall GCN parameters remain $O(Kd_h^2)$. Compared with conventional fully connected attention fusion ($O(Kd_h^2)$ applied to the fusion itself), the computational overhead of the proposed fusion is significantly reduced.

3.4.2. Temporal Modeling and Probabilistic Output

An embedding vector is introduced for each monitoring point to characterize individual differences. Let the embedding dimension be d_e :

$$E = \text{Embedding}(\text{point_ids}) \in \mathbb{R}^{N \times d_e} \quad (24)$$

The fused spatial features are concatenated with the embedding vectors along the feature dimension:

$$H_{concat} = [H_{fused} || E] \in \mathbb{R}^{N \times (d_h + d_e)} \quad (25)$$

The sequence H_{concat} obtained for each time step constitutes the temporal input sequence, which is modeled using a two-layer *LSTM*:

$$h_t, c_t = LSTM(x_t, h_{t-1}, c_{t-1}) \quad (26)$$

where x_t represents the input features at time t , and h_t and c_t denote the hidden state and memory cell, respectively. The *LSTM* gating computations are:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (27)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (28)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (29)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (30)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (31)$$

$$h_t = o_t \odot \tanh(c_t) \quad (32)$$

To perform multi-step forecasting and improve stability over long prediction horizons, an improved parameter-shared encoder–decoder structure is adopted. In the encoding phase, historical sequences are used to update the hidden states, while in the decoding phase, future sequences are recursively generated under shared parameters. Unlike the conventional Seq2Seq framework with separate encoder and decoder networks, this structure is implemented by continuously unfolding the same *LSTM* module across the historical observation period and the future prediction period.

Training Phase (Teacher Forcing): At the t -th decoding step, the ground-truth observation is used as input:

$$\mu_t, \sigma_t, h_t, c_t = Model(y_{t-1}, h_{t-1}, c_{t-1}) \quad (33)$$

Inference Phase (Autoregressive): The predicted mean is used as the input for the next step:

$$\mu_t, \sigma_t, h_t, c_t = Model(\mu_{t-1}, h_{t-1}, c_{t-1}) \quad (34)$$

By sharing the same state-transition mechanism in both phases, the model maintains consistency in temporal modeling and reduces mismatch between the two stages. Using the predicted mean as the input for the next step may lead to underestimation of output variance. However, given the short-term prediction window and NLL-based calibration during training, the practical impact of this potential underestimation on the prediction results is relatively limited.

The *LSTM* hidden states are transformed into probabilistic outputs through two independent linear layers:

$$\mu_t = W_\mu h_t + b_\mu \quad (35)$$

$$\sigma_t = Softplus(W_\sigma h_t + b_\sigma) \quad (36)$$

where the *Softplus* function $Softplus(x) = \log(1 + e^x)$ ensures $\sigma_t > 0$. The model output is assumed to follow an approximately Gaussian distribution:

$$p(y_t | X_{1:t}) = N(y_t; \mu_t, \sigma_t^2) \quad (37)$$

The reasonableness of this assumption is empirically supported by residual diagnostic analyses, including residual autocorrelation tests and distributional shape analysis. Under

extreme events or abnormal operating conditions, the residuals may exhibit more pronounced skewness or heavy-tailed behavior, in which case the Gaussian likelihood may be limited in its ability to characterize abnormal responses.

3.4.3. Training Strategy and Optimization

The model is optimized using the Adam optimizer for parameter updates, with an initial learning rate set to 0.005. An exponential decay learning rate schedule is applied, where the learning rate decays by a factor of 0.95 after each epoch. To prevent overfitting, an early stopping mechanism is introduced, which halts the training process if the validation loss does not improve for 10 consecutive epochs.

To simultaneously optimize prediction accuracy and uncertainty estimation, the model adopts the Gaussian Negative Log-Likelihood (NLL) loss function as the training objective. Let the predicted mean and standard deviation at time t be μ_t and σ_t , respectively; the loss function is:

$$\begin{aligned}\mathcal{L} &= -\sum_{t=1}^T \log p(y_t | \mu_t, \sigma_t) \\ &= \sum_{t=1}^T \left[\log \sigma_t + \frac{(y_t - \mu_t)^2}{2\sigma_t^2} + \frac{1}{2} \log(2\pi) \right]\end{aligned}\quad (38)$$

This loss function optimizes both prediction accuracy (through the squared error term) and uncertainty estimation (through the log-variance term), naturally penalizing overconfident predictions and encouraging appropriate uncertainty calibration.

3.5. Risk Warning and Assessment

3.5.1. Confidence Intervals and Risk Classification

The predicted mean μ and standard deviation σ enable the construction of confidence intervals:

$$CI_{1-\alpha} = [\mu_t - z_{\alpha/2}\sigma_t, \mu_t + z_{\alpha/2}\sigma_t] \quad (39)$$

where $z_{\alpha/2}$ is the critical value of the standard normal distribution. For a 95% confidence interval, $z_{0.025} = 1.96$.

Based on the relationship between predictions and the safety threshold θ , a four-level risk classification system is defined, as shown in Table 1.

Table 1. Four-level risk classification system.

Level	Name	Condition	Action
0	Normal	$\mu < 90\%\theta, \sigma < 30\%\theta$	Routine Monitoring
1	Attention	$\mu \geq 90\%\theta$ or $\sigma \geq 30\%\theta$	Enhanced Monitoring
2	Warning	$\mu + 2\sigma > \theta(95\%CI)$	Notify Relevant Parties
3	Alarm	$\mu > \theta$	Immediate Intervention

Note that the thresholds used in this study ($90\%\theta$ and $30\%\theta$) are not fixed engineering standards. They are a reasonable implementation within the probabilistic prediction framework for risk warning. The risk warning process follows these steps: real-time prediction generates μ and $\sigma \rightarrow$ construct confidence interval $\rightarrow$ calculate deviation degree from safety threshold $\rightarrow$ classify risk level according to predefined criteria $\rightarrow$ trigger warning when threshold is exceeded.

3.5.2. Evaluation Metrics

To evaluate model prediction performance, this study uses four metrics: Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), Normalized Deviation (*ND*), and Coefficient of Determination (R^2). Each metric is defined as follows:

$$RMSE = \sqrt{\frac{1}{NT'} \sum_{n=1}^N \sum_{t=1}^{T'} (\mu_{n,t} - y_{n,t})^2} \quad (40)$$

$$MAE = \frac{1}{NT'} \sum_{n=1}^N \sum_{t=1}^{T'} |\mu_{n,t} - y_{n,t}| \quad (41)$$

$$ND = \frac{\sum_{n=1}^N \sum_{t=1}^{T'} |\mu_{n,t} - y_{n,t}|}{\sum_{n=1}^N \sum_{t=1}^{T'} |y_{n,t}|} \quad (42)$$

$$R^2 = 1 - \frac{\sum_{n=1}^N \sum_{t=1}^{T'} (\mu_{n,t} - y_{n,t})^2}{\sum_{n=1}^N \sum_{t=1}^{T'} (y_{n,t} - \bar{y})^2} \quad (43)$$

Lower *RMSE*, *MAE*, and *ND* values indicate smaller prediction errors. Higher R^2 values indicate better ability to explain data variation.

4. Results and Analysis

4.1. Experimental Setup

This study conducts an experimental analysis based on continuous GNSS monitoring data and synchronized environmental factor data. The data are collected from 23 monitoring stations, recording displacements in the X, Y, and Z directions, resulting in a total of 69 time series. The temporal coverage spans from mid-2023 to the end of 2024, with daily scale continuous observations capable of characterizing short- to medium-term surface deformation. All GNSS sequences are strictly aligned in the temporal dimension, providing reliable data for multivariate time series forecasting.

To characterize the influence of external conditions on surface deformation, three environmental driving variables (water level, temperature, and precipitation) are introduced as exogenous inputs, synchronized with the GNSS data and incorporated into the model training. Additionally, spatial coordinate information corresponding to each GNSS variable is provided to support the construction of spatio-temporal modeling approaches.

The experiments are conducted in a unified software and hardware environment, utilizing an AMD Ryzen 7 5800H processor (Advanced Micro Devices, Inc., Santa Clara, CA, USA) with 16 GB of memory. Model training and prediction are implemented using Python 3.12.7 and PyTorch 2.8.0. The dataset is chronologically partitioned: 70% of the data is used as the training set, 15% as the validation set, and 15% as the test set, ensuring that the model uses only historical data for prediction. A sliding window approach is employed, with an input window of 30 time steps, an output window of 7 time steps, and a stride of 1 day.

To validate the effectiveness of the proposed method, typical models including LSTM, TCN, CNN-LSTM, and STGCN are selected as baseline comparisons. Experiments are conducted under identical data partitioning and training strategies. Model performance is evaluated using metrics such as *RMSE*, *MAE*, *ND*, and R^2 , where *RMSE* and *MAE* measure prediction error, *ND* eliminates scale effects, and R^2 assesses the model's ability to fit GNSS displacement variation trends.

4.2. Multi-Source Spatial Graph Structure Analysis

Figure 3 presents the four types of spatial adjacency graphs constructed in this study, together with their quantitative topological statistics, namely the DTW temporal similarity

graph, Gaussian spatial proximity graph, structural partition graph, and measurement line connectivity graph. These graphs characterize the spatial relationships among GNSS monitoring points from different perspectives, and their distinct topological structures provide complementary information for subsequent multi-graph fusion modeling.

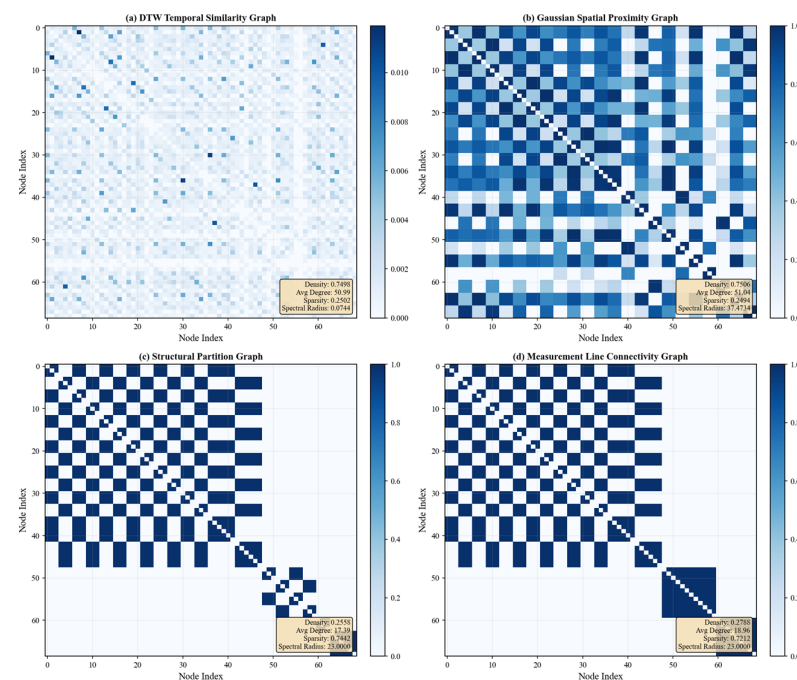


Figure 3. Visualization and topological statistics of multi-source spatial adjacency graphs: (a) DTW temporal similarity graph; (b) Gaussian spatial proximity graph; (c) structural division graph; (d) measurement-line connectivity graph. The insets in each subfigure report the density, average degree, sparsity, and spectral radius of the corresponding graph.

From Figure 3, it can be observed that the DTW temporal similarity graph (a) and the Gaussian spatial proximity graph (b) are both relatively dense weighted graphs overall, with densities of 0.7498 and 0.7506, and average degrees of 50.99 and 51.04, respectively. However, the former exhibits a more dispersed connectivity pattern, reflecting the consistency of temporal evolution among monitoring points, whereas the latter shows more pronounced locally continuous high-weight regions near the diagonal, capturing spatial proximity. In contrast, the structural partition graph (c) and the measurement line connectivity graph (d) are relatively sparse binary graphs, with densities of 0.2558 and 0.2788, respectively, exhibiting block-diagonal and banded connectivity patterns to represent engineering partition consistency and structural constraints among monitoring points located on the same measurement line. Overall, the differences in connectivity patterns and topological characteristics among these four graphs indicate that they can characterize the relationships among monitoring points from different granularities and physical perspectives, thereby jointly supporting multi-graph fusion modeling.

4.3. Prediction Performance Evaluation

4.3.1. Time-Series Prediction Analysis at Representative Monitoring Points

This study presents comparative time-series prediction results for four representative monitoring points (S1, S23, S46, and S69), as illustrated in Figure 4. These monitoring points were selected to represent different spatial locations and deformation patterns. In Figure 4, the solid blue line denotes the predicted mean values, the red dashed line represents the ground-truth observations, and the light blue and dark blue shaded regions correspond to the 95% and 68% confidence intervals, respectively.

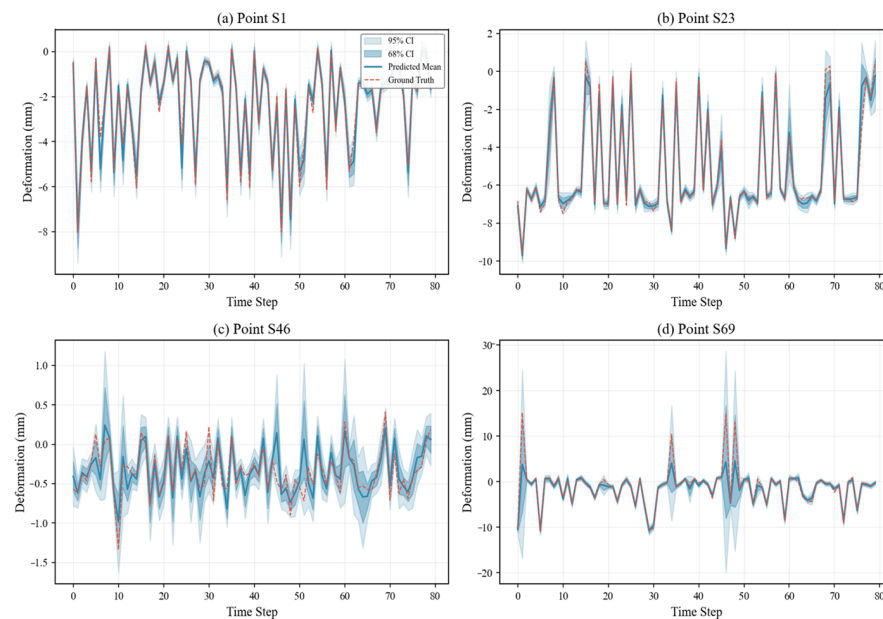


Figure 4. Time series prediction results for four representative monitoring points (S1, S23, S46, and S69). These points were selected to represent different spatial locations, deformation magnitudes, and signal characteristics.

Overall, as shown in Figure 4, the predicted curves exhibit strong agreement with the ground-truth observations across all monitoring points, demonstrating the model's ability to capture periodic variations, local abrupt changes, and deformation patterns of different magnitudes. For monitoring points with relatively large deformation amplitudes (e.g., S1, S23, and S69), the model accurately tracks both peak values and abrupt transition points. In contrast, for monitoring points with smaller deformation amplitudes and more stable signals (e.g., S46), where the deformation ranges approximately from -1.5 mm to 1.5 mm, the predictions remain stable and consistent with the observations. The prediction confidence intervals appropriately widen during periods of pronounced deformation variations and become relatively tighter during stable phases, indicating that the model can effectively characterize uncertainty under different temporal regimes.

Considering the prediction results across all four representative monitoring points, the proposed STMGAFN model demonstrates strong predictive capability and generalization performance across different spatial locations, deformation magnitudes, and deformation patterns. Beyond accurate mean prediction, the model effectively quantifies predictive uncertainty, providing a reliable basis for subsequent risk assessment.

4.3.2. Global Prediction Accuracy Analysis

Figure 5 presents a scatter plot of the predicted values versus the ground-truth observations for all monitoring points over the entire test set, which is used to evaluate the global fitting accuracy of the model. In the figure, the black dashed line represents the ideal prediction line ($y = x$), while the red solid line denotes the fitted regression line. The distribution of the scatter points reflects the overall characteristics of the prediction errors.

As can be observed, the majority of scatter points are tightly clustered around the ideal prediction line, and the fitted regression line closely aligns with the ideal line, indicating that the model exhibits no evident systematic bias in its predictions. Further, the fitted regression line has a slope of 0.9799 and an intercept of -0.0445 mm, indicating good agreement between the predicted and ground-truth values. The model achieves an R^2 value of 0.9865 on the test set, demonstrating strong global fitting capability.

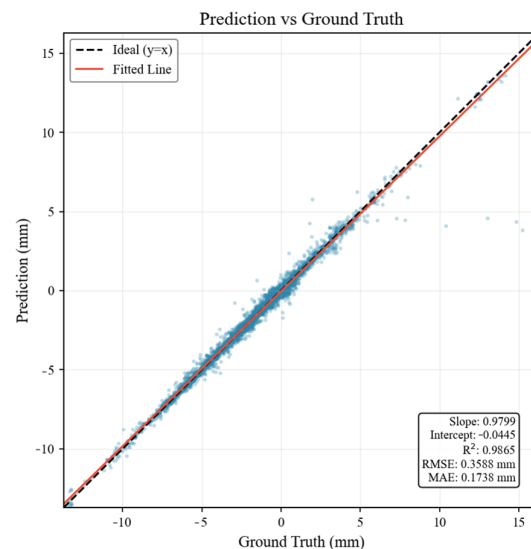


Figure 5. Global prediction accuracy scatter plot for all monitoring points over the entire test set. Each scatter point represents a single prediction sample.

In addition, the prediction errors remain at a low level, with $RMSE = 0.3588$ mm and $MAE = 0.1738$ mm, indicating that the proposed model can deliver stable and high-accuracy predictions across a wide range of deformation magnitudes. Overall, the global prediction accuracy analysis confirms that the proposed STMGAFN model performs effectively in multi-point, long-term dam deformation prediction tasks, maintaining consistently high predictive accuracy under different deformation magnitudes and deformation patterns.

4.3.3. Prediction Error Distribution Characteristics

Figure 6 presents the histogram of prediction errors on the test set along with the fitted normal distribution curve, and further examines the consistency between the error distribution and the theoretical normal distribution using a Q–Q plot. The prediction error is defined as the difference between the observed value and the predicted value, and its distribution characteristics reflect the systematic bias and random error of the model predictions.

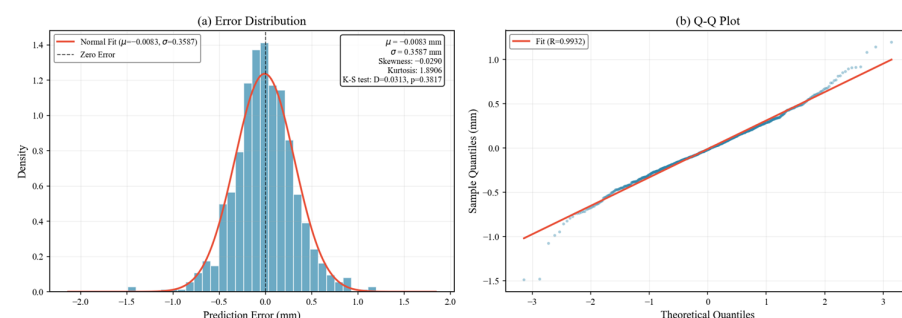


Figure 6. Prediction error distribution analysis. (a) Histogram of prediction errors with a fitted normal distribution curve ($\mu = -0.0083$ mm, $\sigma = 0.3587$ mm); (b) Q–Q plot comparing sample quantiles of prediction errors against theoretical normal quantiles, where alignment with the diagonal indicates approximate normality.

As shown in Figure 6a, the prediction errors exhibit an approximately bell-shaped distribution overall, and the fitted normal curve agrees well with the empirical histogram. The mean error is $\mu = -0.0083$ mm, which is close to zero, indicating that the model does not exhibit a noticeable systematic bias. The standard deviation of the errors is

$\sigma = 0.3587$ mm, reflecting the overall variability of the prediction errors. The Kolmogorov–Smirnov test yields $D = 0.0313$ and $p = 0.3817$, suggesting that the null hypothesis that the errors follow a normal distribution cannot be rejected at the 5% significance level.

In the Q–Q plot shown in Figure 6b, the sample quantiles generally align well with the theoretical normal quantiles, indicating a strong linear relationship. This suggests that the prediction errors can be regarded as approximately normally distributed, providing a reasonable basis for subsequent uncertainty quantification under the Gaussian assumption.

4.3.4. Performance Comparison by Monitoring Point Across Individual Metrics

Figure 7 presents performance comparisons for all monitoring points across four evaluation metrics (ND, RMSE, MAE, R^2), analyzing the impact of spatial heterogeneity on prediction accuracy. Each subplot corresponds to one evaluation metric, with the horizontal axis representing monitoring point numbers, the vertical axis representing metric values, and the red dashed line indicating the average level across all monitoring points.

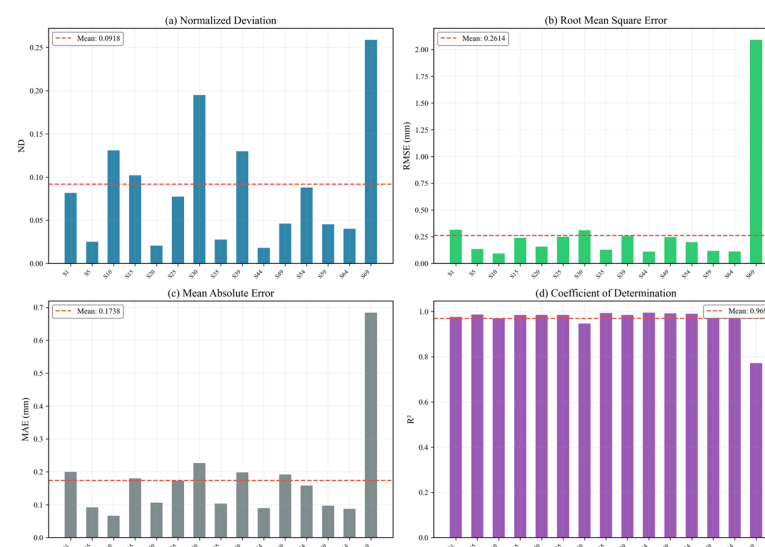


Figure 7. Performance comparison by monitoring point across individual metrics: (a) normalized deviation (ND); (b) root mean square error (RMSE, mm); (c) mean absolute error (MAE, mm); (d) coefficient of determination (R^2). The horizontal axis represents monitoring point indices (S1–S69), and the red dashed line indicates the average value across all monitoring points.

The results demonstrate that the STMGAFN model maintains stable high-precision prediction performance at the vast majority of monitoring points. Most monitoring points exhibit ND values below the average level (0.0918), with $RMSE$ and MAE concentrated in the 0.1–0.3 mm and 0.05–0.2 mm ranges, respectively. The R^2 values exceed 0.95 for nearly all points, with an average of 0.9693, indicating strong capability in fitting deformation trends.

However, monitoring point S69 shows notable deviations, with significantly elevated error metrics ($RMSE \approx 2.0$ mm, $ND \approx 0.26$, $R^2 \approx 0.77$), representing the lowest prediction accuracy among all monitoring points. Data quality diagnostics reveal that S69 (corresponding to the vertical displacement component of station Y40, located in the lower-left zone of the spillway) exhibits a missing rate of 12.1% with a maximum consecutive gap of 38 days, and a signal-to-noise ratio of only 13.19 dB. The relatively high missing rate introduces imputation uncertainty that propagates to downstream prediction, while the low SNR—consistent with the inherently lower measurement precision of GNSS in the vertical direction—increases the difficulty of distinguishing true deformation signals from measurement noise. Additionally, the spillway location of this station results in structural and loading conditions that differ from the main dam body, leading to weaker spatial

correlations with neighboring monitoring points and limiting the effectiveness of spatial information propagation through graph convolution. These factors collectively account for the reduced prediction accuracy at this point, while the overall model performance remains robust across the majority of monitoring points.

4.4. Uncertainty Quantification and Calibration Analysis

4.4.1. Reliability Verification of Uncertainty Estimation

Figure 8a presents the correlation analysis results between predicted uncertainty (standard deviation σ) and actual prediction error ($|\mu - y|$). The horizontal axis represents the model-predicted uncertainty σ , the vertical axis represents the absolute value of actual prediction error, each scatter point represents a prediction sample, and the red fitted line represents the linear relationship between the two.

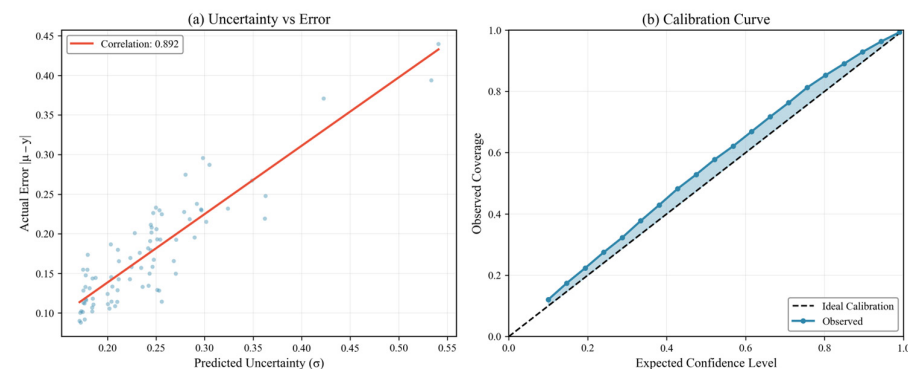


Figure 8. Uncertainty quantification and calibration analysis: (a) scatter plot showing the correlation between model-predicted uncertainty (σ) and absolute prediction error ($|\mu - y|$); (b) probabilistic prediction calibration curve comparing expected confidence levels (horizontal axis) with observed coverage rates (vertical axis).

A strong positive correlation ($r = 0.892$) demonstrates that the model's uncertainty estimates effectively reflect actual prediction error trends, indicating good error awareness capability. Although some scatter dispersion exists, suggesting room for improvement in certain special deformation patterns, the overall results validate the proposed probabilistic prediction mechanism's ability to provide reasonable uncertainty characterization across different prediction samples.

4.4.2. Probabilistic Prediction Calibration Quality Assessment

Figure 8b presents the calibration curve comparing expected confidence levels with observed coverage rates. The black dashed line represents the ideal calibration ($y = x$), while the blue solid line indicates the actual model performance.

In the low confidence level region (0–0.4), the calibration curve almost perfectly coincides with the ideal line, indicating that the model provides accurate probability estimates for relatively certain samples. In the medium confidence level region (0.4–0.8), the calibration curve slightly deviates from the ideal line, but the deviation is minimal, maintaining good calibration performance. In the high confidence level region (0.8–1.0), the calibration curve again converges toward the ideal line, indicating that the model maintains accurate coverage rates when constructing high confidence intervals.

Overall, the model's actual coverage rate in Figure 8b closely follows the ideal calibration line ($y = x$), showing neither significant overconfidence nor excessive conservatism. This indicates that the predicted confidence intervals are well calibrated, validating the reliability of the probabilistic outputs from the proposed STMGAFN model and supporting their practical use in risk warning.

4.4.3. Dynamic Evolution Analysis of Uncertainty at Individual Monitoring Points

Figure 9 takes monitoring point S1 as an example, presenting the temporal evolution characteristics of the model's predicted mean and confidence intervals (upper panel), along with the temporal distribution of corresponding prediction residuals (lower panel), to analyze the dynamic behavior of uncertainty at the individual monitoring point scale.

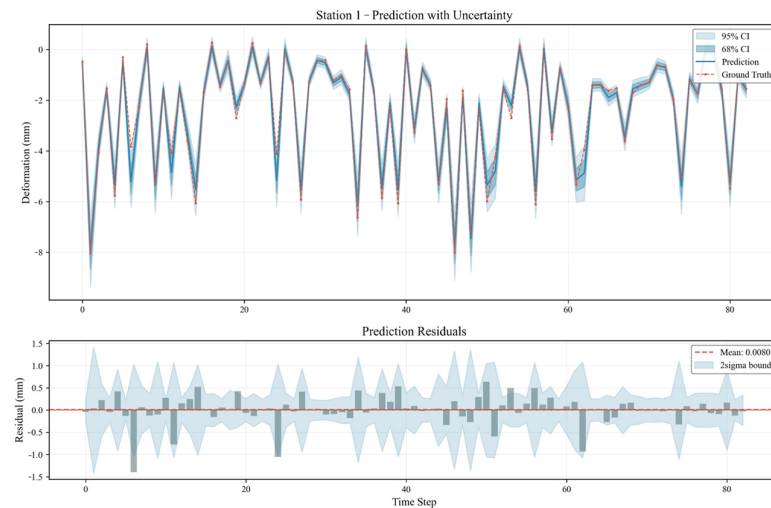


Figure 9. Dynamic evolution analysis of uncertainty at monitoring point S1. (**Upper panel**) predicted mean (blue line) versus ground-truth observations (red dashed line) with 95% and 68% confidence interval bands. (**Lower panel**) temporal distribution of prediction residuals, with the $\pm 2\sigma$ bounds indicated.

The predicted mean maintains strong agreement with the ground-truth observations throughout the entire time series. The prediction confidence intervals adaptively adjust according to deformation states, appropriately expanding during periods of large deformation fluctuations and significantly converging during relatively stable periods, indicating that the model's uncertainty estimation can reflect changes in time series deformation complexity.

Prediction residuals are randomly distributed around zero overall, showing no significant systematic bias or temporal correlation. Residual magnitudes mostly fall within the $\pm 2\sigma$ range, with a mean close to zero, indicating that the model maintains good unbiasedness and well-calibrated uncertainty at the individual monitoring point level. These results demonstrate that the proposed model not only achieves stable deformation prediction at the individual monitoring point level but also reasonably characterizes the dynamic evolution characteristics of uncertainty, with behavior consistent with the aforementioned global statistical analysis results.

4.5. Model Comparison and Ablation Experiments

To systematically evaluate the prediction performance of the proposed STMGAFN model and verify its advantages over typical time-series and spatio-temporal graph models, LSTM, TCN, CNN-LSTM, and STGCN are selected as baseline models, and comparative experiments are conducted under identical data partitioning and evaluation metrics.

Figure 10 presents comprehensive performance comparison results for each model across four metrics: *RMSE*, *MAE*, *ND*, and R^2 .

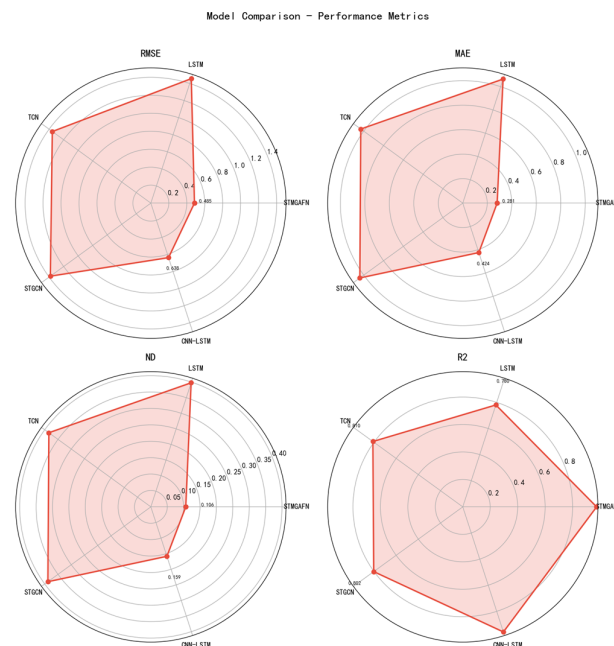


Figure 10. Model performance comparison across four evaluation metrics: $RMSE$ (mm); MAE (mm); ND ; R^2 . Models compared include LSTM, TCN, CNN-LSTM, STGCN, and the proposed STMGA FN. Lower values of $RMSE$, MAE , and ND indicate better performance; higher R^2 indicates better fit.

Overall, the STMGA FN model achieves the best performance across all four evaluation metrics, with significantly lower error values ($RMSE$, MAE , and ND) and higher goodness-of-fit ($R^2 = 0.9865$) than all baseline models. This demonstrates superior prediction accuracy and capability in characterizing deformation evolution trends under different scale conditions.

Among baseline models, STGCN performs relatively well as a single-graph spatio-temporal method but still lags significantly behind STMGA FN, indicating that relying solely on a single spatial topology is insufficient to fully characterize the complex spatial coupling relationships in dam deformation processes. In contrast, STMGA FN's joint modeling and adaptive fusion of multi-source spatial graphs enables capturing correlation features from multiple spatial perspectives, validating the effectiveness of multi-graph fusion and probabilistic modeling strategies in complex spatio-temporal deformation prediction tasks.

To analyze the contribution of each key module in the STMGA FN model to overall performance, a series of ablation experiments are designed, comparing prediction performance changes by progressively removing or simplifying model components. The results are shown in Table 2.

Table 2. Ablation study results comparison.

Model Configuration	$RMSE$	MAE	ND	R^2
Full Model	0.3588	0.1738	0.09723	0.9865
Without Denoise	0.6459	0.4048	0.1509	0.9575
Without Attention	0.5595	0.3819	0.1424	0.9681
Without Distance	0.4674	0.2860	0.1085	0.9770
Only DTW	0.4590	0.3169	0.1202	0.9778
Only Gaussian	0.4766	0.3164	0.12004	0.9760

The ablation study results demonstrate that each key module contributes significantly to the model's performance. Removing the denoising module (Without Denoise) causes the most severe performance degradation, with $RMSE$ increasing from 0.3588 mm to 0.6459 mm

(an 80.0% increase), indicating that data quality has a fundamental impact on prediction accuracy. Eliminating the attention-based fusion mechanism (Without Attention) raises *RMSE* to 0.5595 mm (a 55.9% increase), confirming the critical role of adaptive weighting in multi-graph information fusion. When only a single graph structure is retained (Only DTW: 0.4590 mm; Only Gaussian: 0.4766 mm), performance remains inferior to the multi-graph fusion configuration, highlighting that joint modeling better captures the complex spatial relationships among monitoring points. Removing the spatial distance constraint (Without Distance) also reduces prediction accuracy (*RMSE* increases to 0.4674 mm), demonstrating the irreplaceable role of spatial proximity information in deformation prediction. Overall, the ablation results indicate that the denoising process, multi-graph fusion mechanism, and attention-weighted strategy all play important roles in enhancing the performance of STMGAFN.

4.6. Risk Warning Effectiveness Validation

To validate the effectiveness of the proposed model in risk warning scenarios, this study constructs a four-level risk classification system based on predicted means, uncertainty estimates, and safety thresholds, conducting comprehensive analysis from the perspectives of risk distribution, threshold response, and uncertainty participation in decision-making. The results are shown in Figure 11.

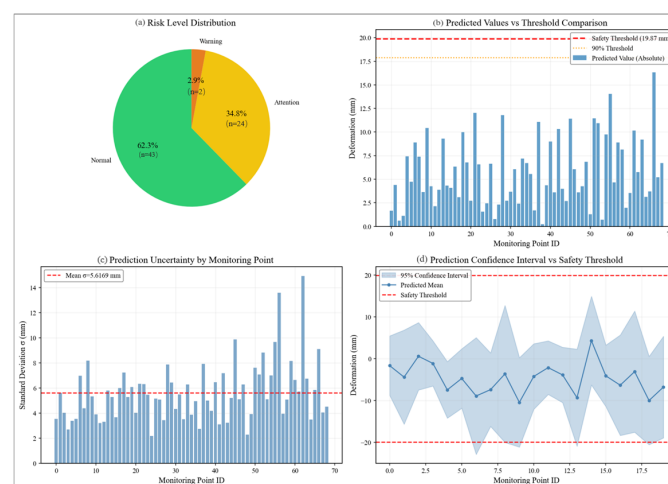


Figure 11. Validation of risk warning effectiveness: (a) distribution of the four risk levels (Normal, Attention, Warning, and Alarm), with absolute counts annotated; (b) comparison between the predicted deformation values at each monitoring point (blue bars) and the safety threshold (red dashed line); (c) distribution of uncertainty (σ) at each monitoring point; (d) relationship between the confidence interval and the safety threshold at representative monitoring points, where the shaded area denotes the 95% confidence interval.

Figure 11a presents the distribution of risk levels. By including the absolute counts for each risk category (Normal: $n = 43$, 62.3%; Attention: $n = 24$, 34.8%; Warning: $n = 2$, 2.9%), the statistical basis of the risk distribution can be more intuitively understood. The high proportion of the Normal category reflects the actual operational status of the dam, where most monitoring periods correspond to safe or low-risk conditions, while high-risk events are inherently rare. Figure 11b demonstrates that the vast majority of monitoring points remain below safety thresholds, with only a few approaching or exceeding warning thresholds, validating the threshold-based risk screening capability.

Figure 11c reveals significant differences in prediction uncertainty across monitoring points, providing supplementary reliability information for risk assessment. Figure 11d illustrates that some monitoring points, despite having predicted means below thresholds,

exhibit confidence intervals that approach or cover the threshold, thereby being reasonably classified into higher risk levels. This reflects the necessity of incorporating uncertainty in risk determination.

The comprehensive results show that the proposed STMGAFN method achieves high prediction accuracy, strong robustness, and reliable uncertainty estimation under real monitoring conditions. These results suggest that the method has strong potential for real-time dam safety monitoring and risk warning. The method can provide quantitative support for identifying abnormal deformation, issuing risk warnings, and allocating inspection and maintenance resources during dam operation. In particular, when the risk level reaches the warning or alarm stage, the prediction results and their confidence intervals can provide a more reliable basis for engineering risk assessment. They can also support timely on-site inspection, intensified monitoring, and preventive intervention. At the same time, the real-time delivery of graded warning information can help relevant authorities take emergency actions in time. It can also support emergency resource allocation and on-site response arrangements. In this way, the method can improve the timeliness, reliability, and rationality of safety monitoring and operational management.

5. Discussion

The proposed Spatio-temporal Multi-Graph Attention Fusion Network (STMGAFN) demonstrates clear advantages in multi-point dam deformation prediction, outperforming baseline models (LSTM, TCN, CNN-LSTM, and STGCN) in terms of prediction accuracy, robustness, and risk warning effectiveness. By integrating multi-source spatial topology modeling, data-quality-aware preprocessing, and probabilistic time-series prediction, the framework provides a unified solution for multi-point deformation prediction under incomplete monitoring conditions. The following sections further discuss key methodological choices, engineering implications, potential limitations, and future research directions.

5.1. Multi-Source Heterogeneous Graph Modeling

The multi-source heterogeneous graph structure characterizes spatial dependencies among monitoring points from several complementary perspectives, including the DTW temporal similarity graph, Gaussian spatial proximity graph, structural partition graph, and measurement-line connectivity graph. Compared with spatio-temporal models relying on a single graph structure, multi-graph fusion provides a more comprehensive representation of spatial coupling relationships in dam structural responses.

Ablation results show that the multi-graph configuration reduces *RMSE* by 21.8% and 24.7%, respectively, compared with models using only the DTW graph or only the Gaussian spatial graph. These results confirm that different spatial topologies provide complementary information for deformation prediction. Meanwhile, the lightweight adaptive fusion mechanism proposed in this study introduces only $O(K)$ learnable parameters, significantly reducing computational overhead compared with conventional fully connected attention-based fusion approaches. In addition, the learned fusion weights are interpretable and directly reflect the relative importance of different spatial graph structures.

It should be noted that the current structural partition graph and measurement-line graph adopt binary adjacency relationships. In future studies, weighted structural partition graphs or weighted measurement-line graphs may be explored. By incorporating spatial distance within partitions, elevation differences, or structural stress characteristics into the adjacency weights, more refined spatial correlations under engineering structural constraints could be modeled.

5.2. Data Quality Assurance Strategy

To address the widespread issues of missing values and noise in monitoring data, this study proposes a hierarchical adaptive preprocessing strategy that improves data quality through three-level missing data imputation and VMD–wavelet joint denoising. Ablation experiments show that removing the denoising module increases *RMSE* from 0.3588 mm to 0.6459 mm, demonstrating that data quality assurance plays a fundamental role in prediction performance.

Compared with deep learning-based imputation methods proposed in recent studies (e.g., GAN-based or multi-head attention imputation models), the two approaches differ in design philosophy. Deep learning-based imputation methods generally possess stronger nonlinear modeling capability and can learn complex data distributions under irregular missing patterns. However, these models often require more complex training procedures and higher computational cost, and they may suffer from training instability in scenarios with limited data or high missing rates. Moreover, purely data-driven imputation approaches often lack mechanisms to explicitly incorporate engineering prior knowledge, such as temporal continuity of deformation sequences and physical spatial correlations among neighboring monitoring points.

In contrast, the hierarchical strategy proposed in this study adaptively selects different imputation methods according to the duration of missing intervals. Short-term gaps are filled using cubic spline interpolation to preserve temporal continuity consistent with the smooth evolution of dam deformation. Long-term gaps are estimated using bidirectional GBR regression with environmental covariates, incorporating both temporal context and physical driving factors. Boundary gaps are addressed using spatial neighborhood weighted interpolation, which maintains spatial correlation structures among monitoring points. This strategy achieves higher computational efficiency than deep generative models, as it avoids adversarial training or complex neural architectures, while maintaining clear physical interpretability. Nevertheless, for larger datasets or more complex missing patterns (e.g., systematic missing segments), deep learning-based imputation methods may offer advantages. Future research may explore hybrid frameworks that combine physics-informed hierarchical strategies with deep learning-based imputation techniques.

5.3. Uncertainty Quantification and Distribution Assumptions

In this study, probabilistic prediction is implemented using a Gaussian negative log-likelihood loss, enabling the model to jointly estimate prediction mean and uncertainty. Experimental results indicate a strong correlation between predicted uncertainty and actual prediction errors, suggesting that the model effectively reflects the reliability of prediction results.

However, the current framework assumes that prediction errors follow a Gaussian distribution. While this assumption is reasonable under normal operational conditions, error distributions may exhibit significant skewness or heavy-tailed behavior during extreme deformation events or abnormal operating conditions. In such cases, the Gaussian assumption may underestimate the probability of extreme deviations. In addition, during inference the model adopts an autoregressive prediction strategy, where the predicted mean is used as the input for the next step. This mechanism may theoretically lead to slight underestimation of predictive variance. Although the effect is limited under the current short-term prediction horizon (7 days), it may become more pronounced for longer forecasting horizons.

Future research may improve uncertainty modeling through several directions:

(1) Introducing more flexible distribution assumptions, such as Student's *t*-distribution to better capture heavy-tailed error behavior or mixture density networks (MDN) for multi-

modal uncertainty modeling; (2) incorporating lightweight Monte Carlo (MC) sampling calibration strategies, where a small number of MC samples are introduced during inference to correct variance underestimation while controlling additional computational cost, making the approach suitable for real-time engineering monitoring systems.

5.4. Risk Classification Threshold Design

The risk classification thresholds used in this study ($90\%\theta$ and $30\%\theta$) represent a practical implementation within the probabilistic prediction framework, aiming to integrate prediction mean and uncertainty information into risk decision-making. These thresholds are not calibrated based on historical anomaly events but are instead defined as an initial configuration based on general principles of engineering safety management and characteristics of probabilistic prediction.

In practical engineering deployment, these threshold parameters should be dynamically adjusted according to dam-specific safety regulations, historical operational experience, and accumulated warning–response records. As monitoring data continue to accumulate, especially when events approaching or exceeding safety thresholds are observed, statistical analysis can be conducted to further calibrate threshold parameters, thereby achieving a better balance between early-warning sensitivity and false alarm rates.

5.5. Computational Efficiency and Scalability

From an engineering deployment perspective, computational efficiency is an important consideration. Under the current experimental configuration (23 GNSS monitoring stations and 69 time series), the model can be efficiently trained on a standard workstation environment (AMD Ryzen 7 5800H CPU with 16 GB memory). The lightweight multi-graph fusion mechanism and parameter-sharing temporal modeling structure effectively control overall model complexity.

However, when monitoring networks expand to larger scales (e.g., more than 100 monitoring nodes), computational complexity may increase significantly, particularly due to dense graph structures and the DTW similarity matrix calculation. Future research may enhance scalability through several strategies: applying graph sparsification techniques (e.g., k-nearest-neighbor pruning) to reduce graph convolution computation; adopting hierarchical graph modeling or subgraph training strategies to handle large-scale monitoring networks; and introducing sampling-based graph convolution methods such as GraphSAGE to support efficient training with large node numbers.

In addition, because the current dataset mainly reflects normal operational states of the dam, the model's performance under extreme deformation events or abnormal operating conditions requires further investigation. Under extreme conditions, deformation processes may exhibit stronger non-stationarity or responses outside the training distribution. Future work may construct simulated extreme-condition datasets or augmented datasets to systematically evaluate model robustness under abnormal scenarios.

5.6. Summary and Future Perspectives

Overall, the STMGAFN framework provides a promising intelligent modeling approach for dam deformation prediction and risk warning under complex monitoring conditions. By integrating multi-source spatial topology modeling, data-quality-aware preprocessing, and probabilistic prediction mechanisms, the proposed method achieves both high prediction accuracy and reliable uncertainty representation.

Future research will focus on the following directions: (1) exploring adaptive graph learning methods and weighted engineering topology graphs to further enhance spatial modeling capability; (2) introducing more flexible probabilistic distribution models and MC sampling calibration strategies to improve uncertainty reliability under extreme conditions;

(3) integrating real-time meteorological forecast data (e.g., precipitation and temperature predictions) as forward-looking covariates, rather than relying solely on historical environmental observations, to enhance the model's anticipatory capability for deformation events driven by extreme weather conditions; (4) conducting cross-dam transfer learning studies to evaluate model generalization across different engineering environments; (5) integrating multi-modal monitoring data sources (e.g., InSAR and leveling observations) to further expand the applicability of the proposed framework.

6. Conclusions

This study focuses on several key problems in dam deformation prediction. These problems include complex spatial relationships among multiple monitoring points, unstable monitoring data quality, and the need for uncertainty information in risk warning. To deal with these problems, this study proposes a Spatio-temporal Multi-Graph Attention Fusion Network (STMGAFN). The proposed method achieves high-accuracy deformation prediction and probabilistic risk warning, even under incomplete data conditions.

The results demonstrate that by introducing a data-quality-aware hierarchical adaptive imputation mechanism and a VMD-wavelet joint denoising strategy, the model maintains stable prediction performance under data quality degradation conditions. Ablation experiments show that removing the denoising module increases the *RMSE* from 0.3588 mm to 0.6459 mm, confirming the fundamental role of data quality assurance in prediction accuracy. Building upon this foundation, the constructed multi-source spatial topology modeling framework characterizes spatial correlation relationships among monitoring points from four complementary dimensions—temporal similarity, spatial proximity, structural partitioning, and measurement line connectivity—significantly enhancing the model's expressive capability for complex deformation patterns. The lightweight adaptive attention fusion mechanism achieves effective integration of multi-graph features while reducing computational complexity, and the multi-graph fusion configuration yields notable performance improvements over single-graph configurations.

On real dam monitoring data comprising 23 GNSS stations and 69 time series, STMGAFN achieves a prediction accuracy of $RMSE = 0.3588$ mm, $MAE = 0.1738$ mm, and $R^2 = 0.9865$, outperforming baseline models including LSTM, TCN, CNN-LSTM, and STGCN. Through parameter-shared recursive probabilistic temporal modeling, the model jointly outputs deformation means and prediction uncertainties, with a correlation coefficient of 0.892 between predicted uncertainty and actual error, and a calibration curve closely aligned with the ideal line. The four-level risk classification system constructed based on confidence intervals forms a closed-loop decision support mechanism of “prediction → quantification → classification → warning”. Nevertheless, the current framework has certain limitations, including the reliance on predefined graph structures, the potential insufficiency of the Gaussian distribution assumption under extreme deformation events, the lack of historical anomaly event validation for the risk classification system, and the scalability to large-scale monitoring networks (>100 nodes) that remains to be further evaluated.

Future work will focus on the following directions: (1) exploring adaptive graph learning methods and weighted engineering topology graphs to further enhance spatial modeling capability; (2) introducing more flexible probabilistic distribution models (e.g., Student's *t*-distribution or mixture density networks) and lightweight Monte Carlo sampling calibration strategies to improve uncertainty reliability under extreme conditions; (3) integrating real-time meteorological forecast data (e.g., precipitation and temperature predictions) as forward-looking covariates, rather than relying solely on historical environmental observations, to enhance the model's anticipatory capability for deformation

events driven by extreme weather conditions; (4) incorporating multi-platform collaborative monitoring data, such as UAV-based visual inspection and InSAR surface deformation measurements, alongside the existing GNSS observations, to further expand the spatial coverage and information richness of the monitoring input; (5) conducting cross-dam transfer learning and federated learning studies to evaluate model generalization across different engineering environments while preserving data privacy constraints inherent in critical infrastructure monitoring.

In summary, the STMGAFN method proposed in this study provides an intelligent solution for dam deformation prediction and risk warning that balances accuracy, stability, and uncertainty expression, holding significant engineering application value for enhancing the scientific rigor and intelligence level of hydraulic engineering safety monitoring.

Author Contributions: Conceptualization, D.W. and J.L.; methodology, D.W.; software, D.W.; validation, D.W., W.L. and G.Z.; formal analysis, D.W. and W.L.; investigation, D.W. and W.L.; resources, Z.W.; data curation, D.W., Z.W. and G.Z.; writing—original draft preparation, D.W.; writing—review and editing, D.W., J.L. and N.X.; visualization, D.W. and W.L.; supervision, N.X. and J.L.; project administration, N.X.; funding acquisition, N.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Anhui Provincial Water Conservancy Science and Technology Project, grant number slky202501-06; by the National Natural Science Foundation of China (NSFC), grant number 61971178; in part by the Ordos Science and Technology Major “Open Bidding for Selecting the Best Candidates” Project, grant number JBGS-2023-002; in part by the Anhui Province Key Laboratory of Intelligent Building & Building Energy Saving (Anhui Jianzhu University), grant number IBES2025KF02ZD; and by the State Grid Corporation of China Headquarters Technology Project, grant numbers 5500-202140127A and 5500-202440171A-1-1-ZN.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to thank all colleagues and staff who provided technical and administrative support during this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

To systematically evaluate the effect of the VMD mode number K on denoising performance, VMD-wavelet joint denoising experiments were conducted for $K = 2$ to $K = 8$. The results were analyzed from four perspectives: reconstruction error, signal-to-noise ratio (SNR), smoothness ratio, and modal energy distribution. The smoothness ratio is defined as the ratio of the variance of the second-order difference in the denoised signal to that of the original signal and is used to reflect the level of high-frequency roughness retained after denoising. The results are summarized in Table A1.

Table A1. Sensitivity analysis of the VMD mode number K on denoising performance.

K	<i>Recon RMSE (mm)</i>	<i>SNR (dB)</i>	<i>Smoothness Ratio</i>	<i>Min Model Energy (%)</i>
2	0.3213	10.90	0.009	1.15
3	0.2832	12.35	0.125	0.80
4	0.2513	13.60	0.243	0.66
5	0.2166	15.22	0.301	0.62
6	0.1896	16.46	0.404	0.57
7	0.1698	17.53	0.480	0.48
8	0.1444	18.95	0.542	0.49

Based on the results in Table A1, this study ultimately adopts $K = 3$. This choice was not made solely by minimizing reconstruction error, but rather as a compromise based

on the sensitivity analysis over $K = 2 \rightarrow 8$, jointly considering denoising strength, signal fidelity, modal energy distribution, and physical interpretability. Specifically, when $K = 2$, the smoothness ratio is only 0.009, indicating overly aggressive denoising that may suppress meaningful medium-frequency deformation information; when K increases to 3, the smoothness ratio rises markedly to 0.125, suggesting that the modal decomposition begins to more effectively distinguish noise components from meaningful signal structures. As K continues to increase, although the reconstruction error and SNR keep improving, the change in smoothness ratio becomes more gradual. At the same time, a larger K leads to greater mode splitting, increased model complexity, and reduced compactness and interpretability of the decomposition results. From a physical perspective, $K = 3$ corresponds well to the three dominant components commonly present in GNSS displacement signals, namely the low-frequency trend, the medium-frequency periodic response, and the high-frequency noise. Therefore, it helps preserve physical interpretability while avoiding modal redundancy caused by over-decomposition. In addition, we further examined the effect of different K values on the final prediction performance, and the results show that the model performance varies only slightly when K is changed within a reasonable range, indicating that the overall effectiveness of the denoising module is more critical than the exact value of K . Consistently, the clear performance degradation observed in the ablation setting “Without Denoise” further confirms the necessity of the denoising module as a whole. Overall, $K = 3$ was selected not because it is absolutely optimal for every metric, but because it provides a good balance among denoising effectiveness, signal fidelity, physical interpretability, and computational efficiency under the current data and task setting.

Appendix B

This study performed normality tests on the finest-scale wavelet coefficients of the highest-frequency mode after VMD, and provides representative histogram fitting plots (Figure A1) and Q–Q plots (Figure A2).

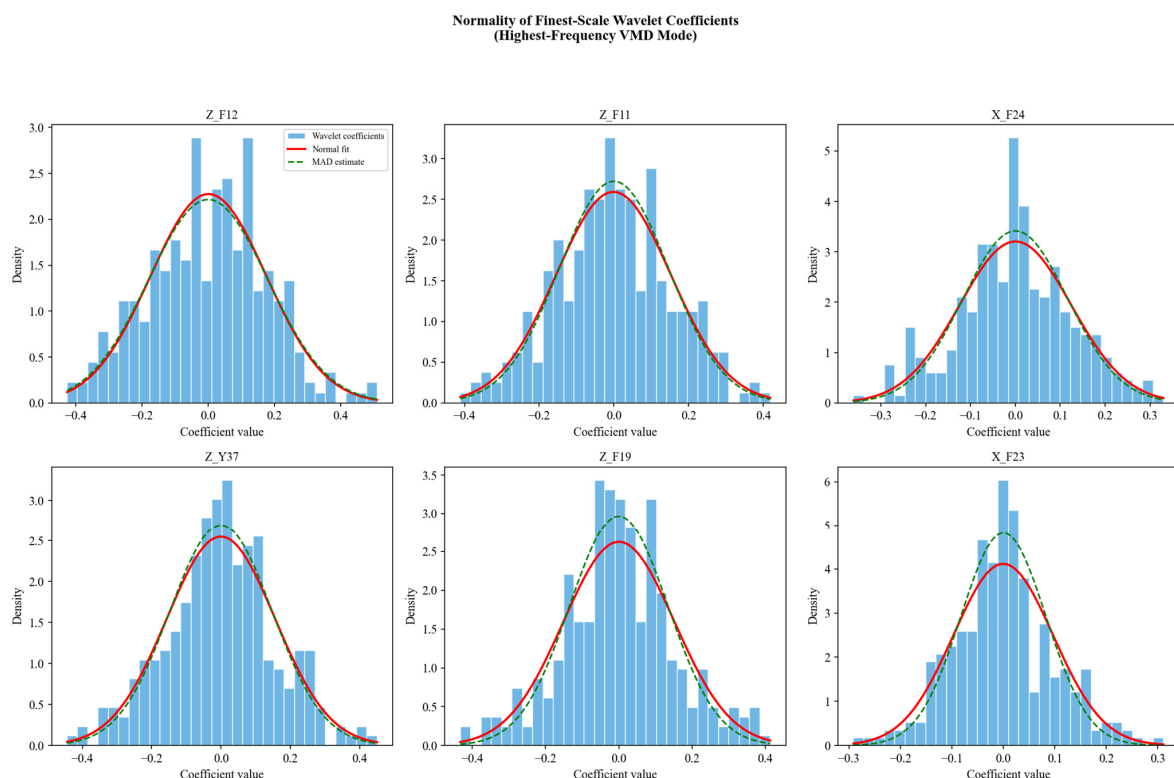


Figure A1. Normality of the Finest–Scale Wavelet Coefficients.

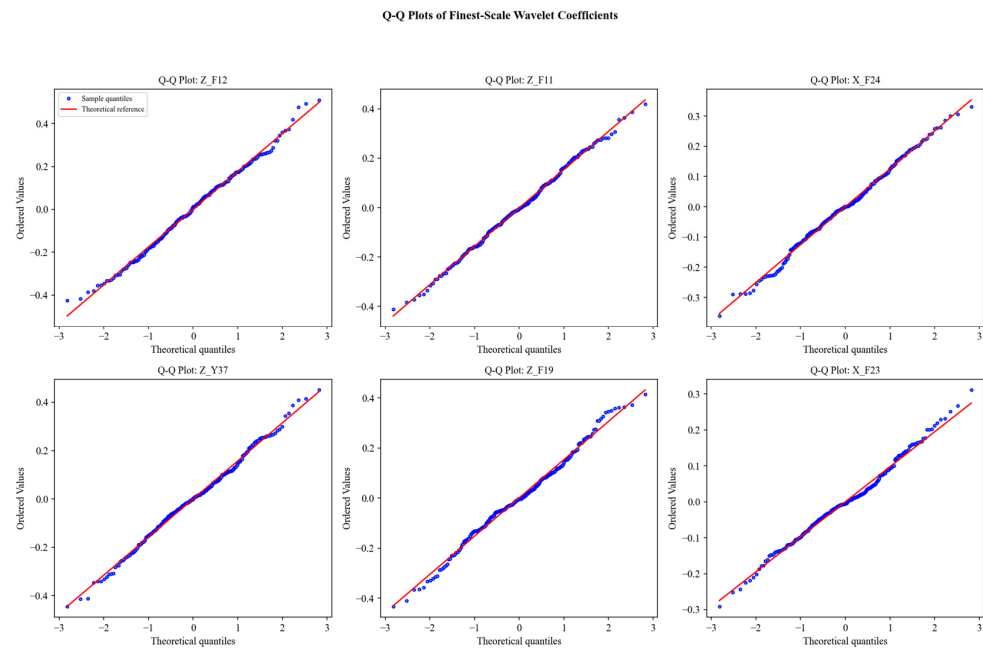


Figure A2. Q-Q Plots of the Finest-Scale Wavelet Coefficients.

The results indicate that these wavelet coefficients exhibit good approximate normality overall: the skewness is close to zero ($|skewness| < 0.18$), the kurtosis is also close to zero ($|kurtosis| < 0.4$), and the p -values of all three normality tests are well above the significance level of 0.05 (see Table A2 for details). Correspondingly, the data points in the Q–Q plots lie mostly close to the theoretical reference line, suggesting that the universal threshold is empirically applicable in this pre-separated setting.

Table A2. Normality Test Results for Wavelet Coefficients.

Series	Skewness	Kurtosis	Shapiro p -Value	JB p -Value	D'Agostino p -Value
Z_F12	0.022	−0.195	0.471	0.786	0.842
Z_F11	−0.008	−0.189	0.861	0.805	0.863
X_F24	−0.057	−0.031	0.434	0.919	0.917
Z_Y37	−0.015	0.171	0.527	0.834	0.738
Z_F19	0.011	0.281	0.095	0.620	0.546
X_F23	0.171	0.371	0.110	0.215	0.195

Meanwhile, the noise level was estimated using MAD because of its strong robustness to outliers and mild departures from the assumed distribution. Even when the noise does not strictly follow the ideal Gaussian assumption, MAD still provides a stable scale estimate. Therefore, compared with the direct use of the standard deviation, MAD is more appropriate for threshold calculation in practical monitoring data.

Appendix C

The threshold θ_{DTW} is not a fixed hyperparameter manually tuned by hand; instead, it is determined in a data-driven manner based on the distribution of pairwise DTW similarities among monitoring sequences. Specifically, we first compute the DTW distance between every pair of monitoring sequences and use its reciprocal, $1/DTW$, as the similarity measure. Then, the 25th percentile of all nonzero similarity values is selected as the threshold θ_{DTW} . To evaluate the sensitivity of this threshold choice, we further conducted a systematic analysis over the 10th to 90th percentile range and assessed the resulting graph structures from three aspects: graph density, average degree, and spectral gap. The results are summarized as follows (Table A3):

Table A3. Graph Properties vs. DTW Threshold Percentiles.

Percentile	θ_{DTW}	Number of Edges	Graph Density	Average Degree	Spectral Gap
10%	0.000348	2111	0.900	61.2	0.0097
25%	0.000473	1759	0.750	51.0	0.0051
50%	0.000717	1173	0.500	34.0	0.0016
75%	0.001225	587	0.250	17.0	0.0000
90%	0.002018	235	0.100	6.8	0.0000

Based on the table above, the 25th percentile is ultimately chosen as θ_{DTW} . The main reason is that this threshold provides a good balance between graph connectivity and the filtering of weak connections. Compared to the overly dense graph structure corresponding to the 10th percentile (graph density 0.90), the 25th percentile effectively filters out some weakly similar edges, reducing the interference of noisy connections on feature propagation. Compared to the sparse graphs corresponding to the 75th and 90th percentiles, the 25th percentile still maintains good global connectivity. In other words, the selection of θ_{DTW} is determined jointly by the distribution of the data and the sensitivity analysis of the graph structure, rather than being a fixed heuristic.

References

1. Cao, X.; Sheng, J.; Jiang, C.; Yuan, D.; Zhang, H. Concrete dam deformation prediction model considering the time delay of monitoring variables. *Sci. Rep.* **2025**, *15*, 8458. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Wang, Q.; Gao, Y.; Gong, T.; Liu, T.; Sui, Z.; Fan, J.; Wang, Z. Dam surface deformation monitoring and analysis based on PS-InSAR technology: A case study of xiaolangdi reservoir dam in China. *Water* **2023**, *15*, 3298. [\[CrossRef\]](#)
3. Song, X.; Deng, L.; Wang, H.; Zhang, Y.; He, Y.; Cao, W. Deep learning-based time series forecasting. *Artif. Intell. Rev.* **2024**, *58*, 23. [\[CrossRef\]](#)
4. Ren, Q.; Li, M.; Fang, X.; Xiao, L.; Li, H. A separate modeling approach to noisy displacement prediction of concrete dams via improved deep learning with frequency division. *Adv. Eng. Inform.* **2024**, *60*, 102367. [\[CrossRef\]](#)
5. Liu, X.; Li, Z.; Sun, L.; Khailah, E.Y.; Wang, J.; Lu, W. A critical review of statistical model of dam monitoring data. *J. Build. Eng.* **2023**, *80*, 108106. [\[CrossRef\]](#)
6. Zhou, Y.; Bao, T.; Li, G.; Shu, X.; Li, Y. Multi-expert attention network for long-term dam displacement prediction. *Adv. Eng. Inform.* **2023**, *57*, 102060. [\[CrossRef\]](#)
7. Lei, W.; Wang, J.; Ji, T.; Li, P. Dam deformation early warning model based on cluster analysis and spatiotemporal data fusion. *Measurement* **2022**, *204*, 112109. [\[CrossRef\]](#)
8. Sahili, Z.A.; Awad, M. Spatio-temporal graph neural networks: A survey. *arXiv* **2023**, arXiv:2301.10569. [\[CrossRef\]](#)
9. Ren, Q.; Li, M.; Chen, Y.; Ji, X.; Liu, H. Efficient prediction uncertainty quantification in dam behavior monitoring with attention-based sequence-to-sequence learning. *Appl. Soft Comput.* **2024**, *167*, 112321. [\[CrossRef\]](#)
10. Xu, R.; Liu, X.; Wei, J.; Ai, X.; Li, Z.; He, H. Predicting the deformation of a concrete dam using an integration of long short-term memory networks and Kolmogorov–Arnold networks with a dual-stage attention mechanism. *Water* **2024**, *16*, 3043. [\[CrossRef\]](#)
11. Cai, S.; Gao, H.; Zhang, J.; Peng, M. A self-attention-LSTM method for dam deformation prediction based on CEEMDAN optimization. *Appl. Soft Comput.* **2024**, *159*, 111615. [\[CrossRef\]](#)
12. Zhou, B.; Wang, Z.; Fu, S.; Chen, D.; Yin, T.; Gao, L.; Zhao, D.; Ou, B. Dam deformation prediction model based on multi-scale adaptive kernel ensemble. *Water* **2024**, *16*, 1766. [\[CrossRef\]](#)
13. Liu, M.; Wen, Z.; Su, H. Deformation prediction based on denoising techniques and ensemble learning algorithms for concrete dams. *Expert Syst. Appl.* **2024**, *238*, 122022. [\[CrossRef\]](#)
14. Wei, H.; Liu, X.; Wang, F.; Ai, X. An integrated deep learning model for predicting concrete dam deformation with multi-point spatiotemporal correlation. *Measurement* **2025**, *256*, 118546. [\[CrossRef\]](#)
15. Cao, E.; Bao, T.; Yuan, R.; Hu, S. Hierarchical prediction of dam deformation based on hybrid temporal network and load-oriented residual correction. *Eng. Struct.* **2024**, *308*, 117949. [\[CrossRef\]](#)
16. He, P.; Pan, J.; Li, Y. Long-term dam behavior prediction with deep learning on graphs. *J. Comput. Des. Eng.* **2022**, *9*, 1230–1245. [\[CrossRef\]](#)
17. Lu, T.; Gu, H.; Gu, C.; Shao, C.; Yuan, D. A multi-point dam deformation prediction model based on spatiotemporal graph convolutional network. *Eng. Appl. Artif. Intell.* **2025**, *149*, 110483. [\[CrossRef\]](#)

18. Liu, J.; Li, Q.; Li, L.; An, S. Structural damage detection and localization via an unsupervised anomaly detection method. *Reliab. Eng. Syst. Saf.* **2024**, *252*, 110465. [[CrossRef](#)]
19. Liu, J.; Li, Q.; Li, L.; An, S. Unsupervised structural damage detection and severity assessment via U-GraphFormer. *Struct. Health Monit.* **2025**, 14759217251348421. [[CrossRef](#)]
20. Wang, X.; Bao, T.; Gong, J.; Zhang, J.; Xu, Y. Missing data imputation method for dams based on multi-head attention mechanisms and generative adversarial networks. *Eng. Struct.* **2025**, *336*, 120350. [[CrossRef](#)]
21. Xiao, S.; Yang, J.; Cheng, L.; Li, G.; Jia, D.; Chai, J. Feature imputation and label prediction of dam deformation behavior based on Graph Representation Learning. *KSCE J. Civ. Eng.* **2026**, *30*, 100381. [[CrossRef](#)]
22. Ahmad, H.; Zhang, Y.; Rehman, H.; Alam, M.; Ullah, Z.; Shahid, M.A.; Khan, M.; Siddique, A. A hybrid spatio-temporal graph attention (ST D-GAT framework) for imputing missing SBAS-InSAR deformation values to strengthen landslide monitoring. *Remote Sens.* **2025**, *17*, 2613. [[CrossRef](#)]
23. Wang, Q.A.; Lu, A.W.; Ni, Y.Q.; Wang, J.F.; Ma, Z.G. Bayesian Network in Structural Health Monitoring: Theoretical Background and Applications Review. *Sensors* **2025**, *25*, 3577. [[CrossRef](#)]
24. Sillionis, N.E.; Anyfantis, K.N. Data-driven probabilistic quantification and assessment of the prediction error model in damage detection applications. *Probabilistic Eng. Mech.* **2023**, *71*, 103412. [[CrossRef](#)]
25. Kong, X.; Chen, Z.; Liu, W.; Ning, K.; Zhang, L.; Marier, S.M.; Liu, Y.; Chen, Y.; Xia, F. Deep learning for time series forecasting: A survey. *Int. J. Mach. Learn. Cybern.* **2025**, *16*, 5079–5112. [[CrossRef](#)]
26. Wang, Z.; Ma, G.; Ai, Z.; Ding, Q.; Zhou, W. Noise adaptive spatiotemporal neural networks for deformation prediction of high rockfill dams. *Expert Syst. Appl.* **2025**, *294*, 128836. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.