Structures 57 (2023) 105277

Contents lists available at ScienceDirect

Structures

journal homepage: www.elsevier.com/locate/structures

Enhanced data imputation framework for bridge health monitoring using Wasserstein generative adversarial networks with gradient penalty

Shuai Gao [a], Chunfeng Wan [a,*], Zhenwei Zhou [b], Jiale Hou [c], Liyu Xie [d], Songtao Xue [d,e,*]

[a] Southeast University, Key Laboratory of Concrete and Prestressed Concrete Structure of Ministry of Education, Nanjing 210096, China
[b] School of Civil Engineering and Architecture, East China Jiao Tong University, Nanchang 330013, China
[c] School of Civil Engineering, Tsinghua University, Beijing 100084, China
[d] Research Institute of Structural Engineering and Disaster Reduction, College of Civil Engineering, Tongji University, Shanghai 200092, China
[e] Department of Architecture, Tohoku Institute of Technology, Sendai, Miyagi 982-8577, Japan

ARTICLE INFO

Keywords:
Structural health monitoring
Missing data imputation
Generative adversarial networks
Gradient penalty

ABSTRACT

The availability of complete data is essential for accurately assessing structural stability and condition in structural health monitoring (SHM) systems. Unfortunately, data missing is a common occurrence in daily monitoring operations, which hinders real-time analysis and evaluation of structural conditions. Although considerable research has been conducted to efficiently recover missing data, the implementation of these recovery methods often encounters issues such as serious mode collapse and gradient vanishing. To address these challenges, this paper proposes a missing data imputation framework called WGAIN-GP based on Wasserstein Generative Adversarial Network with Gradient Penalty. This framework aims to enhance the stability and convergence rate of the network during the missing data recovery process. The effectiveness and robustness of the proposed method are extensively evaluated using measured acceleration data from a long-span highway-railway dual-purpose bridge. The results of the implementation demonstrate that the proposed method achieves superior recovery performance even under various missing data conditions, including high missing rates of up to 90%. Furthermore, the generality of the method is validated by successfully recovering data from different missing sensors. Additionally, the recovered data is utilized for modal analysis of the bridge's structural state, further verifying the reliability of the recovery method. The proposed recovery method offers several advantages, with its stability and robustness being particularly noteworthy. By significantly enhancing the reliability of the recovered data, this method contributes to improving the overall accuracy and effectiveness of structural health monitoring systems.

1. Introduction

In recent years, the safety and reliability of large infrastructures have become paramount concerns, garnering significant attention from society [1–5]. To effectively assess and monitor the health condition of such infrastructures in real-time, the deployment of structural health monitoring systems (SHMSs) has become crucial, particularly for large-scale structures [6]. The evaluation of structural behavior relies heavily on the integrity and quality of the extensive data provided by these SHMSs. Unfortunately, the occurrence of continuous and random data missing poses an inevitable challenge in SHMSs due to factors like transmitting interference, communication outages, and sensor faults. This missing data can significantly impact the accuracy of data analysis techniques, such as wavelet transform, and severely distort the probability distribution of the data. Consequently, data recovery plays a vital role in ensuring the integrity of the structural analysis process, enabling the identification of unstable and abnormal states of structures.

Over the past few decades, extensive research has been undertaken to address the challenge of missing data recovery in the context of structural assessment. Generally, two prominent categories, namely model-driven and data-driven methods, have emerged as popular approaches for tackling this issue. While some FEM-based methods have been proposed to recover missing data, their effectiveness is hindered by the demanding requirement of highly accurate models, limiting their practical applicability. Subsequently, data-driven methods have gained prevalence in the field, primarily due to their high efficiency in

* Corresponding authors.
E-mail addresses: wan@seu.edu.cn (C. Wan), xue@tongji.edu.cn (S. Xue).

https://doi.org/10.1016/j.istruc.2023.105277
Received 26 July 2023; Received in revised form 18 September 2023; Accepted 19 September 2023
Available online 26 September 2023

recovering missing data across various scenarios. These methods have proven to be effective in addressing the challenges associated with missing data, offering notable advantages over the model-driven approaches.

Currently, data-driven recovery methods can be broadly categorized into two main types: discriminative methods and generative methods. These two methods adopt different analysis approaches. Discriminative models aim to establish a discriminative function with limited samples and directly focus on prediction models without considering sample generation. Shallow discriminative models, such as MissForest and Matrix Completion [7]., have been proposed for missing data recovery. However, their limited ability to handle large datasets hampers their practical application. With the rapid advancement of deep learning, deep discriminative models, including Convolutional Neural Networks (CNN) [8], deep Stacking Networks (DSN) [9], Recurrent Neural Networks (RNN), and Long Short-Term Memory networks (LSTM) [10], have gained popularity in the field of missing data recovery. However, complete datasets are essential for training discriminative models. Similarly, generative models can be classified into shallow generative models and deep generative models. The fundamental idea is to establish a joint probability density model of samples, infer the posterior probability, and then generate the model. Generative models capture the data distribution from a statistical perspective, enabling the reflection of data similarity. Generative recovery methods are widely used nowadays due to their faster learning convergence. Shallow generative models, such as Expectation-Maximization (EM) [11], Denoising Autoencoders (DAE) [12], and multiple imputations by Chained Equations (MICE) [13], have been proposed. Deep generative models incorporate techniques like Restricted Boltzmann Machine (RBM) [14], Autoencoders (AE), Deep Belief Networks (DBN) [15], Sum-Product Networks (SPN) [16], Bayesian networks [17], and Hidden Markov Models (HMM) [18].

In the field of data recovery, three main types of methods can be identified: statistical methods, machine learning methods, and deep learning methods. Statistical methods, such as mean imputation, multiple imputation, and Maximum Likelihood and least-squares estimation with polynomial equations, have been widely used to recover missing data in various structures and have achieved considerable success [19]. However, statistical learning methods have limitations when it comes to recovering continuous missing data.

In recent decades, machine learning has emerged as a promising approach to address the aforementioned challenges. For instance, Ni et al. [20] proposed a recovery method based on Back-Propagation Neural Network (BPNN) and generalized Regression Neural Network (GRNN). Li et al. [21] utilized multi-task Gaussian Process Regression (mGPR) to reconstruct displacement and temperature data affected by ambient and load excitation. Nevertheless, the method is primarily suitable for recovering short-term data located at the beginning or end of the dataset. Additionally, Bayesian Dynamic Linear Models (BDLM) [12], Log-Quantile-Density (LQD), and Reproducing Kernel Hilbert Space (RKHS) methods [22] have been demonstrated to effectively recover lost bridge monitoring data. Ren et al.[23] further extended Bayesian and Tensor analysis techniques to reconstruct random missing strain and temperature data of concrete bridges by representing the missing data as second or third-order tensors and extracting reliable underlying characteristics. In addition, Wan et al applied Bayesian learning method to explore the suitable methodology for structural health monitoring data recovery [24–26].

In recent years, deep learning-based methods have gained significant attention for missing data recovery. For example, Li et al. [27] employed Empirical Mode Decomposition (EMD) and Long Short-Term Memory networks (LSTM) to recover acceleration data from cable-stayed bridges by extracting correlation relationships among the data. However, the proposed recovery model's stability and accuracy were limited by the absence of noise-free vibration responses. To address this limitation, Xia et al. [28] utilized Convolutional Generative Adversarial Networks to reconstruct signals by learning features from low to high frequencies. Furthermore, Fan et al. [29] demonstrated the effectiveness of Densenets for recovering acceleration data from the Guangzhou tower. Additionally, the SegGan architecture, incorporating skip connections and dense blocks, was proposed to reconstruct acceleration data in numerical simulations and steel frames [30]. In the context of continuous missing strain data, Jiang et al. [31] utilized Generative Adversarial Networks (GANs) for reconstruction. Hou [32] introduced an advanced imputation method based on GANs and data augmentation for multi-sensor missing states in SHMSs. Moreover, Jiang [33] proposed the use of U-NET, integrating advanced techniques such as dense connections, skip connections, residual connections, and a perceptual loss function to effectively capture data loss patterns. Chen et al. [34] proposed a hybrid deep-learning and autoregressive model with attention mechanism (DL-AR-ATT) framework to accurately reconstruct structural responses considering data correlations. While recent deep learning-based methods have achieved notable success in recovering missing data for large-scale and long-term datasets, challenges related to model stability still persist.

Generative Adversarial Network (GAN), as a prominent deep generative model, finds widespread applications in various fields [35]. However, the conventional GAN algorithm has limitations in loss calculation during training and interpretation. To meet specific application requirements, many researchers have made substantial improvements to the classic GAN. For example, Wasserstein GAN with weight clipping and gradient penalty has been proposed to enhance the state-of-the-art performance [36]. Ian Goodfellow initially introduced a Generative Adversarial Network with JS Divergence [37], while Arjovsky [38] proposed Wasserstein GAN with Wasserstein distance to improve model stability. However, these approaches are not perfect, as weight clipping leads to issues with enforcing a Lipschitz constraint, resulting in problems such as unstable training, slow convergence, and vanishing gradients. To address these limitations, Gulrajani [39] proposed Wasserstein GAN with Gradient Penalty (WGAIN-GP) to enforce the Lipschitz constraint effectively, mitigating the problems of mode collapse and vanishing gradients. Leveraging the powerful generative capabilities of GAN, several scholars have developed data missing imputation frameworks based on GAN models [31,32,40,41]. However, mode collapse and gradient vanishing issues, similar to those encountered in the original GAN, have hampered the stability and effectiveness of missing data imputation methods.

In this study, a missing data imputation framework is proposed based on Wasserstein GAN with Gradient Penalty (WGAIN-GP) to enhance the robustness and efficiency. The gradient penalty mechanism is introduced to address the challenges of mode collapse and gradient vanishing. This framework not only captures features from the spatiotemporal correlations of sensors but also leverages non-missing data to improve the accuracy of the imputation model. Furthermore, the proposed missing data imputation method has been utilized on a highway-railway dual-purpose bridge to verify high efficiency and effectiveness. The generality of the proposed method is also demonstrated on different missing sensors, and the recovered data are utilized for operational modal analysis.

2. Methodology

2.1. Problem definition

In classic GAN, generator and discriminator are two vital constitutions, and the desirable results are achieved by multiple adversarial of two segments. The input of generator is a random distribution, and the output is an arbitrary distribution. The inputs of discriminator include the real distribution and the generated distribution by the generator, and the outputs are the scores, with the high score for the real distribution and a low score for the distribution generated by generator. In adversarial process, firstly, the parameters of generator are fixed and the parameters of discriminator are updated; then, the parameters of

discriminator are fixed and generator are updated. And so on, the adversarial process will be stopped until the divergence between the output distribution of generator and the real distribution is minimized and discriminator gives a high score to the generated distribution [37]. The distribution divergence is toughly calculated in GAN, and can be expressed on equation (1) and (2). Numerous scholars introduced different divergence to settle out the difficulty of calculation, namely JS divergence and KL divergence, etc. [42]. However, the calculation process is rather complicated, and have difficulties in integrating to calculate the distance. In addition, the overlapping area of the distribution of P_G and P_{data} is relatively small, so JS Divergence cannot meet the requirement to update the better result. Moreover, it is possible to achieve high accuracy with poor results. Therefore, a new distance (Wasserstein distance) is deliberately considered to measure this distance, and the formulation is shown in equation (3). The Wasserstein distance has a tough limitation (1-Lipschitz) yet, and the training of discriminator must be forced to smooth $D(x)$ from $-\infty$ to ∞. Later, gradient penalty is proposed to improve the training and interpretation process, which keep the gradient between P_G and P_{data} being close to 1.

$$G^* = \arg\min_G Div(P_G, P_{data}) \tag{1}$$

$$V(G,D) = E_{y\sim P_{data}}[\log D(y)\uparrow] + E_{y\sim P_G}[\log(1-D(y))\downarrow] \tag{2}$$

$$\max_{D\in 1-Lipschitz}\{E_{x\sim P_{data}}[D(x)]\uparrow - E_{x\sim P_G}[D(x)]\downarrow\} \tag{3}$$

2.2. Detailed configuration of the network

In this paper, to overcome the mode collapse and gradient vanishing in the network training process, Wasserstein Generative adversarial network with gradient penalty is introduced to recover missing data in the structural health system (SHM). It absorbs the generative ability of classical GAN and the stable capability of gradient penalty[43]. GAN is a typical generative model to generate the missing data depending on data correlation and self-creativity. The configuration of the network is shown in Fig. 1. The backbone network contains two main blocks, the generator network, and the critic network. The generator is designed as a generative network to generate the missing data corresponding to real data. The critic is similar to the classifier, being identical to the discriminator of classic GAN. The data from the generator and real data are inputted in critic to distinguish them as real and fake. The proposed architecture is derived from classical GAN and general form of GAIN [43–44]. However, the architecture of WGAIN-GP is evidently more compact, both generator and discriminator are modeled as two fully connected neural nets. These significantly improve the implemented time and feature extraction capability of the network, making the network more efficient without sacrificing accuracy.

The purpose of WGAIN-GP is to estimate unobserved values in each dataset. Based on the available data, the generator produces imputed values, which include both the missing and non-missing parts. After generating these imputed values, the discriminator assesses them as either 'real' or 'fake.' The imputation process continues only when the imputed data closely resembles 'real' data or when the network identifies the imputed data as 'real.'

Improving the accuracy of imputed values for missing parts involves comparing the imputed values for non-missing parts with the original true values to determine if they are a 'real' match. However, for the missing parts (with values as 0), there are no original reference values available for assessment ('real' or 'fake'). In such cases, the discriminator can evaluate whether the imputed values for non-missing parts are 'real' or not.

If the discriminator judges the non-missing part's imputed data as 'fake,' it prompts the generator to continue generating imputed values. This process repeats multiple times until the discriminator determines that the imputed data is 'real.' Conversely, if the discriminator determines that the imputed data is 'real,' the imputation process for all data concludes, retaining only the imputed values for the missing parts, while discarding the imputed values for the non-missing parts. The original data values are preserved.

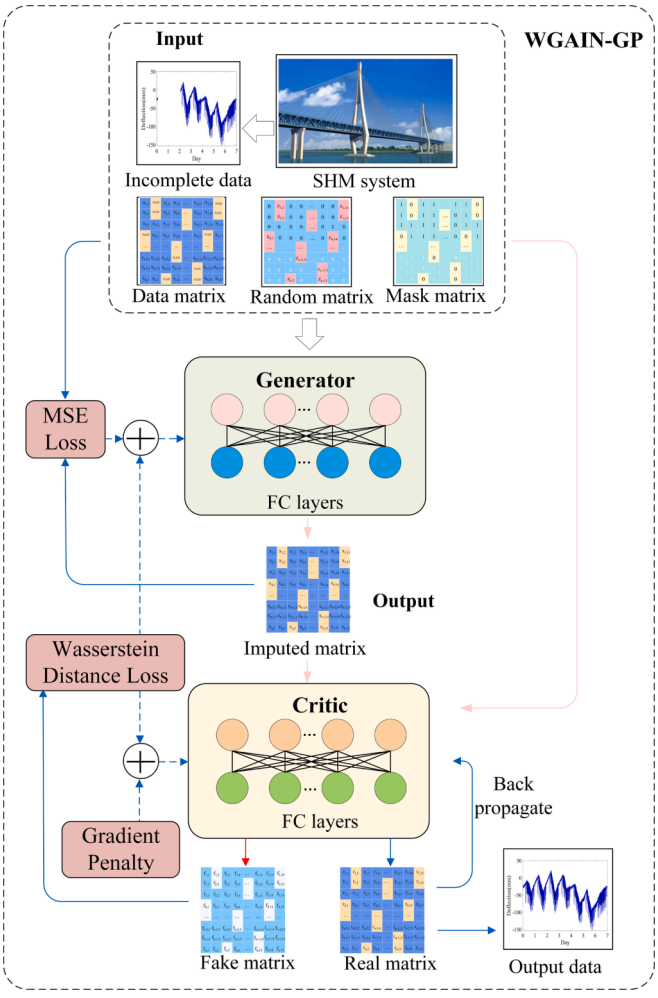

Fig. 1. Detailed configuration of the network.

2.3. Generator

As the critical element in the WGAIN-GP framework, the generator used in the missing data generation is shown in Fig. 2. The incomplete data matrix, generated mask matrix and noise matrix with simple distribution are imported to the generator, which is demonstrated on equation (4). In addition, in an adversarial process, the missing data are recovered with compact network and the imputed data are deemed to the exported element in generator, which is denoted as equation (5). After the training and inference, the final generated data matrix is composed of non-missing data and the estimated missing data, which can be expressed in equation (6). This means that the network is much easier to be trained than the traditional GAN. The activation functions are selected to decide whether a neuron should be activated or not, and the Rectified Linear Unit (ReLU) activation function and hyperbolic tangent (Tanh) activation function are separately utilized to activate the two fully connected layers. The former can be regarded as a non-linear function or piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It is simpler and more effective in the common application, making the network easier to be trained, which is mathematically formed in equation (7). The latter (tanh) is also expressed as sigmoidal (s-shaped) and is shown in equation

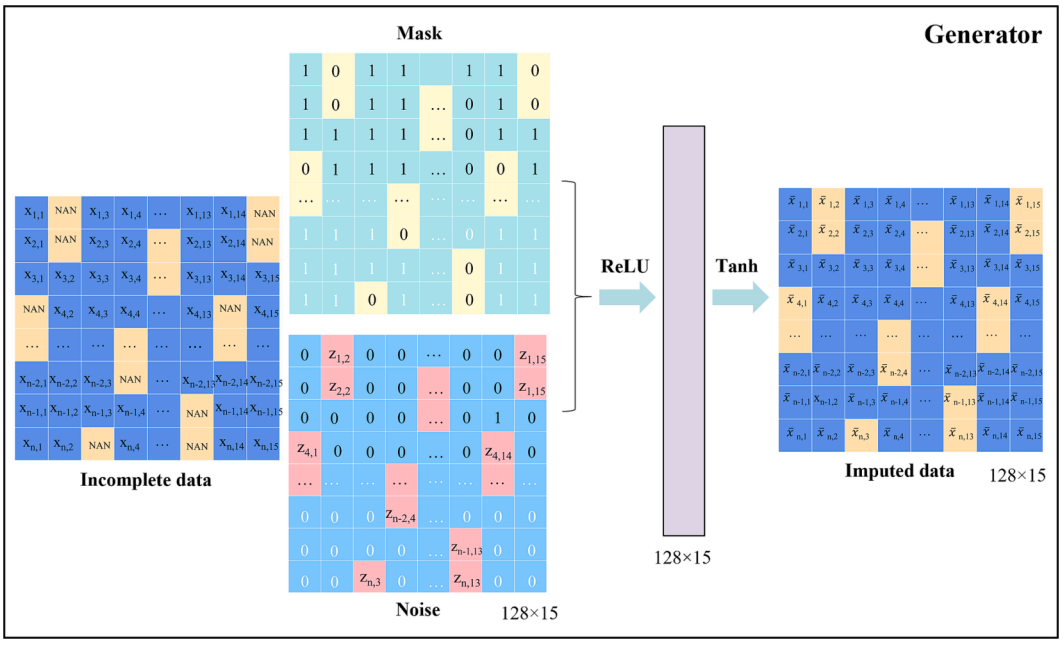

Fig. 2. The architecture of the generator network.

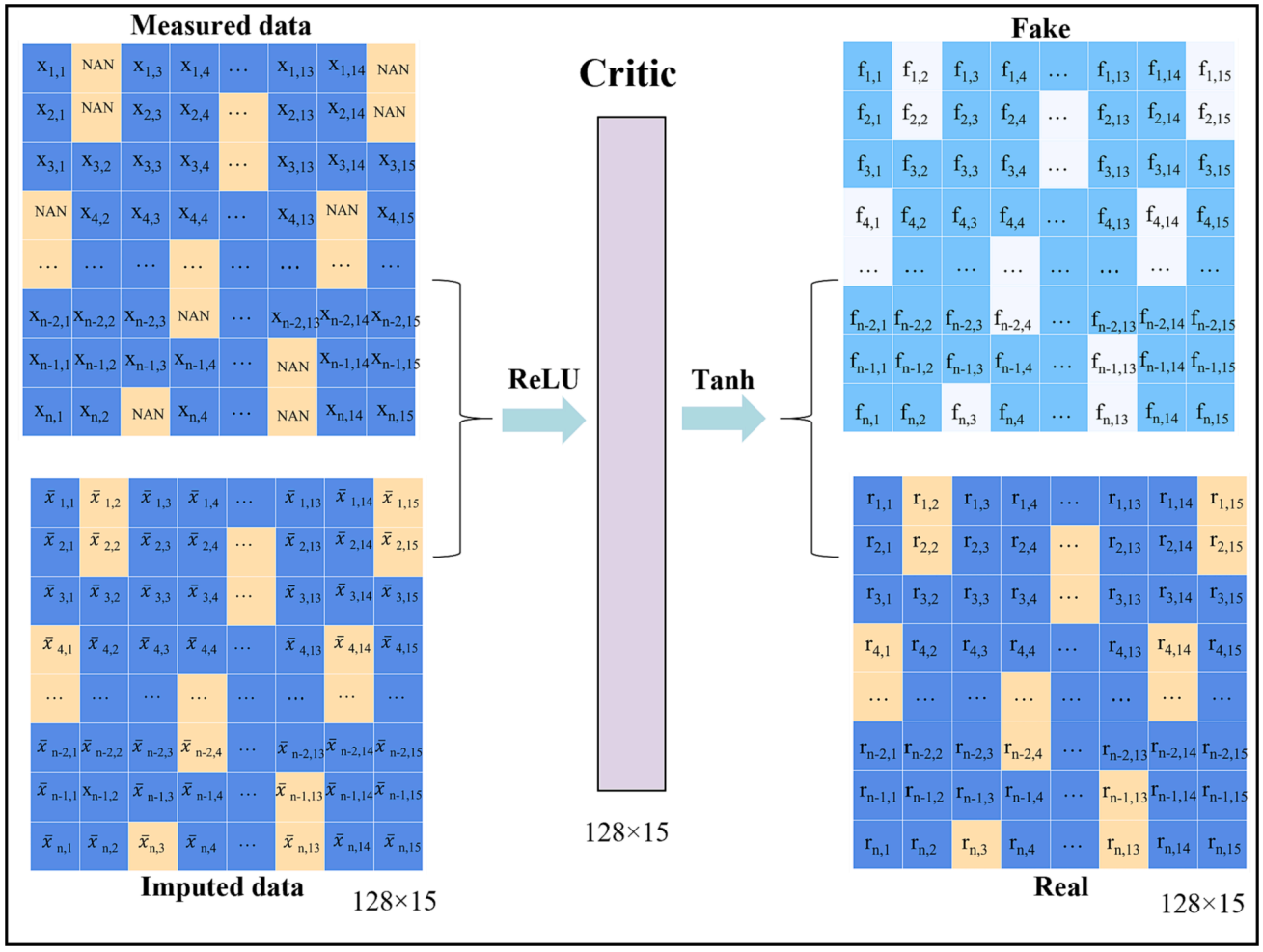

Fig. 3. The architecture of the critic network.

(8). The tanh in output layers is applied to improve the efficiency and accuracy of the missing data recovery. The advantage is that the negative inputs will be mapped strongly negative, and the zero inputs will be mapped near zero in the tanh. The function is differentiable and monotonic while its derivative is not monotonic.

$$\mathbf{Z} = \mathbf{M} \odot \mathbf{X} \oplus (1 - \mathbf{M}) \odot \mathbf{N} \tag{4}$$

$$\bar{\mathbf{X}} = G(\mathbf{Z}, \mathbf{M}) \tag{5}$$

$$\widehat{\mathbf{X}} = \mathbf{M} \odot \mathbf{X} \oplus (1 - \mathbf{M}) \odot \bar{\mathbf{X}} \tag{6}$$

Where $\widehat{\mathbf{X}}$ denotes the data generated by the generator, composing the value of the non-missing part and the imputed value of the missing part; $\bar{\mathbf{X}}$ denotes the imputed data matrix, $\mathbf{X}$ is the incomplete matrix, $\mathbf{Z}$ represents random matrix, $\mathbf{M}$ represents mask matrix, $\mathbf{N}$ is random value of uniform distribution and $\odot$ represents element-wise multiplication.

$$f(x) = \max(0, x) \tag{7}$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{8}$$

2.4. Critic

As another crucial element in the WGAIN-GP framework, we introduce a critic $\boldsymbol{C}$ to distinguish the generated data from generator whether real or fake. In this paper, critic is constituted with two fully connected neural nets, which is foundationally the same as generator. This significantly improve the network's architecture, making the network slimmer and tight. The detailed architecture in critic is shown in Fig. 3. The input of critic is incomplete measured data and the imputation data, and the input is batched with 128 points in data length. In distinguished process, the distance of measured data and generated data in non-missing part are measured in detail. Subsequently, the real data and fake data are outputted in the critic, and the final real data are extracted as unique output step by step.

2.5. Gradient penalty and clipping penalty

To meet the requirement of the Lipschitz constraint, weight clipping is introduced as the implementation of the weight penalty. However, there are a few serious problems with the implementation of weight clipping. The critic loss is designed to maximize the difference between the scores of true and false samples, however, weight clipping independently limits the range of values of each network parameter. Consequently, the optimal strategy is to make all parameters as extreme as possible, either taking the maximum (e.g. 0.01) or the minimum (e.g. −0.01). In addition, weight clipping presumably leads to accidentally disappearing the gradient vanishing or gradient exploding. The reason is that the critic is a multi-layer network, if the clipping threshold is set as a little smaller. After each layer of the network, the gradient becomes a little smaller, multi-layer will be exponentially decaying. Conversely, if set a little larger, after each layer of the network, the gradient becomes a little larger, multi-layer will exponentially explode. Only by setting it not too big or not too small can the generator get just the right back-propagation gradient, however in practice, this balance region can be very narrow. This impairs both the generative and stabilization capabilities of the GAN model.

In this paper, the gradient penalty is introduced to resolve the model collapse and gradient vanishing. After using the gradient penalty, the distribution of the parameter values is much more reasonable after the critic is also fully trained, and the critic can take full advantage of its model fitting ability. In contrast, the gradient penalty can keep the gradient smooth in the backward propagation process. Gradient penalty only applies to the region of true and false samples, as well as the transition zone between the two. However, since it directly limits the gradient norm of the discriminator to around 1, the gradient is very controllable and can be easily adjusted to the appropriate scale size. Gradient penalty can significantly improve the training speed and solve the problem of slow convergence of GAN.

2.6. Loss function and evaluation criteria

In this paper, the innovated missing data imputation with Wasserstein distance and gradient penalty is proposed to alleviate the problem of gradient vanishing/exploding in missing data imputation.

The objective of generator is to minimize the divergence between the measured data and the data generated by generator. Moreover, the generator estimates the whole dataset of missing data. The loss function of the generator is composed of Wasserstein distance loss and mean square error (MSE) loss, which is formed in equation (9). In view of the stable capability of gradient penalty (GP), it can be seen in equation (10) that the loss of critic is attributed to Wasserstein distance loss and gradient penalty. In addition, Adam is regarded as the common Momentum optimizer with Simple operation and efficient calculation. Therefore, the Adam optimizer is selected to update the parameter and gradient descent.

$$\min L_{WGAIN-GP}(\alpha_G) = -E_{\widehat{\mathbf{X}},\mathbf{M}}[(1 - \mathbf{M}) \odot \mathbf{C}(\widehat{\mathbf{X}})] + \alpha E\left[(\mathbf{M} \odot \mathbf{X} - \mathbf{M} \odot \mathbf{G})^2\right] \tag{9}$$

$$\max L_{WGAIN-GP}(\alpha_C) = \underbrace{E_{\widehat{\mathbf{X}},\mathbf{M}}[\mathbf{M} \odot \mathbf{C}(\widehat{\mathbf{X}})] - E_{\widehat{\mathbf{X}},\mathbf{M}}[(1 - \mathbf{M}) \odot (1 - \mathbf{C}(\widehat{\mathbf{X}}))]}_{\text{Wasserstein Distance Loss}} + \underbrace{\lambda\left(\|\nabla_{\hat{x}} C(\mathbf{X}_z)\|_2 - 1\right)^2}_{\text{Gradient Penalty}}$$

where $\widehat{\mathbf{X}}$ denotes the data generated by the generator, composing the value of the non-missing part and the imputed value of the missing part; $\mathbf{X}$ is the incomplete matrix; $\mathbf{X}_z$ denotes the distribution of uniform $\mathbf{P}_{\mathbf{X}_z}$ sampling from the non-missing part of the original part and the imputed part created using the generator. $\mathbf{M}$ represents mask matrix; α represents hyperparameter; and λ represents hyperparameter.

3. Real application and verifications

3.1. Overview of the structure applied

The Tongling Bridge, located at Yangtze River in Anhui Province, China, is a long-span highway-railway dual-purpose bridge. Fig. 4 provides an overview map of the Tongling Bridge. To evaluate the structural condition of the bridge in a quantitative manner, a modern Structural Health Monitoring (SHM) system has been deployed. The design of the SHM system takes into account the significance of the monitoring object and the vulnerability analysis of monitoring positions. One of the crucial monitoring aspects is the vibration of the bridge, which holds immense importance for the safety and overall health of the structure. Hence, monitoring the bridge's acceleration serves as an essential and intuitive parameter to assess the vibration response. In Fig. 4, the layout of the bridge's SHM system consists of 15 acceleration sensors represented by red points. These sensors capture the vibration response induced by vehicle load and ambient load, with a sampling frequency of 100 Hz.

3.2. Dataset preparation

The proposed network is applied to a dataset consisting of 15 acceleration sensors that capture vibrations induced by vehicle load and ambient load, with a sampling frequency of 100 Hz. Among these sensors, seven sensors (J2, J4, J6, J9, J11, J13, and J15) have incomplete data and are selected as the target for data imputation, while the remaining sensors are used for training the network. Data normalization is a crucial step that enhances the efficiency, stability, and accuracy of

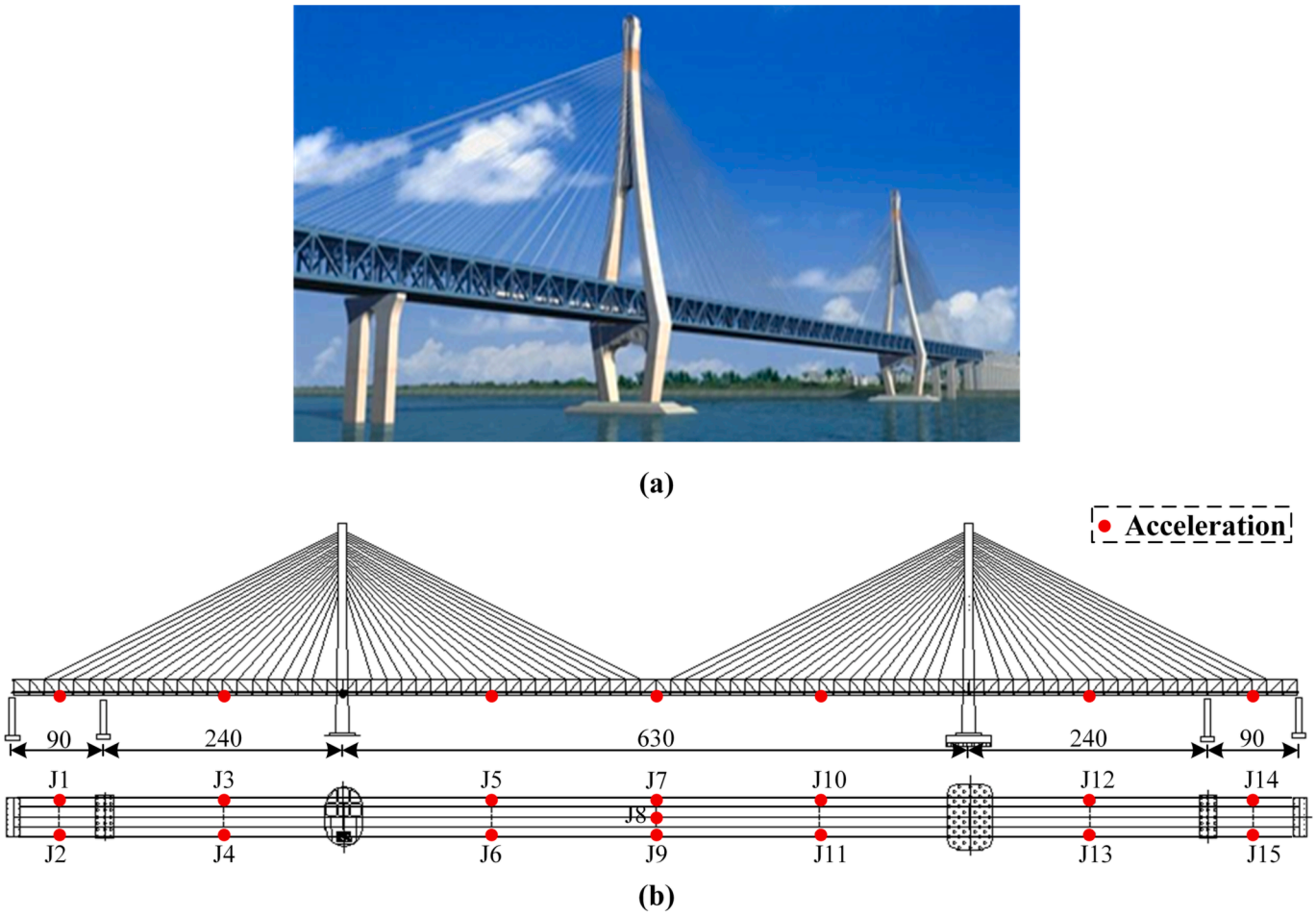

Fig. 4. The schematic chart of the Tongling bridge.

the proposed method while allowing for the analysis of statistical distribution characteristics. The normalization process is performed as follows:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

where x_{norm} is the normalization of the data, and the interval of the normalization is between 0 and 1.

3.3. Model training and hyperparameter setting

The sensor missing data imputation framework was implemented using TensorFlow, and the configurations of the computational platform are two Intel Xeon(R) E5-2696 v4 CPUs, a 256 GB memory, and an NVIDIA TITAN X (Pascal) GPU for boosting algorithm application. The implemented process of the data imputation method will be discussed in detail below.

The hyperparameters of network were extremely important in training process. In this paper, the number of samples in mini-batch was set as constant size (128). The Adaptive moment estimation (Adam) was kept as constant optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. In addition, optimizer's learning rate was 0.0009. Furthermore, $\lambda = 10$ was introduced as a hyper-parameter to compute the critic's loss, and $\alpha = 100$ was used to compute the generator's loss. A number of training iterations were set to achieve the effectiveness and robustness of proposed method.

4. Performance metrics

To qualitatively and quantitatively assess the performance of the proposed missing data imputation model using visual criteria, four performance metrics, namely correlation coefficients (R2), root mean square error (RMSE), mean absolute error (MAE), and Accuracy, are introduced to evaluate the accuracy and robustness of imputation results. These performance metrics can be formulated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{'} - y_i)^2}{\sum_{i=1}^{n}(\bar{y} - y_i)^2} \tag{12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i^{'} - y_i| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^{'} - y_i)^2} \tag{14}$$

$$Accuracy = 1 - \frac{\|y_i - y_i^{'}\|_2^2}{\|y_i^{'}\|_2^2} = 1 - \frac{\sqrt{\sum_{i=1}^{n}(y_i - y_i^{'})^2}}{\sqrt{\sum_{i=1}^{n} y_i^2}} \tag{15}$$

where y_i denotes the measure values of i th location, $y_i^{'}$ represents the imputation values of i th location, $\bar{y}$ demonstrates the mean of the imputation values.

5. Results and discussion

Bridge health monitoring systems commonly encounter two representative forms of missing data: random and continuous. The proposed imputation method efficiently and accurately leverages valuable information in random missing scenarios. Although it may face difficulties in recovering continuous missing data, the proposed method still demonstrates efficient and robust performance even in cases of high proportions of missing data. Moreover, it is important to note that continuous and random data missing can occur simultaneously in the daily operation of SHM systems, a factor often overlooked in missing data recovery research.

This paper introduces three types of missing data—random missing, continuous missing, and hybrid missing (simultaneous random and continuous missing)—to evaluate the generality of the proposed imputation method. Specifically, ten seconds of data from seven incomplete sensors (J2, J4, J6, J9, J11, J13, and J15) are selected to showcase the imputation capacity in both the time and frequency domains.

Frequency domain analysis plays a vital role in detecting structural damage and is a key method for analyzing structural characteristics. In this study, Fast Fourier Transform (FFT) is utilized to transform the signals from the time domain to the frequency domain. Signal processing is performed on the acceleration data, with a sampling frequency of 100 Hz, ensuring that it is more than twice the natural frequency of 50 Hz.

5.1. The generative ability of proposed models

The proposed WGAIN-GP model is employed to recover the missing data of bridge accelerations, showcasing its capacity in various scenarios. For analysis and verification, we focus on data from seven incomplete sensors (J2, J4, J6, J9, J11, J13, and J15) to demonstrate the imputation capabilities in different scenarios, including random missing (Type I), continuous missing (Type II), and hybrid missing (Type III, i.e., random and continuous missing simultaneously).

We first examine the performance of the WGAIN-GP model in the random missing scenario. Fig. 5 illustrates the results of bridge acceleration imputation using the proposed method under a 90% missing rate. It is evident that the WGAIN-GP model exhibits remarkable recovery capacity. This can be attributed to the model's ability to learn spatial–temporal correlations between missing and non-missing data, thus enables effective imputation.

Furthermore, continuous data missing in the sensors is inevitable due to the occasional instability of responses induced by the environment and vehicles. As shown in Fig. 6, the evaluation of continuous missing data for the seven sensors demonstrates the excellent imputation ability of the proposed framework. Overall, the proposed method exhibits superior transferability across different missing sensor scenarios. These results indicate the robustness and efficiency of the proposed framework in imputation performance.

In addition, it is important to note that random and continuous missing can occur simultaneously in bridge SHM systems, and this hybrid missing scenario is often overlooked or challenging to identify. Similarly, as shown in Fig. 7, we utilize seven acceleration sensors to evaluate the imputation ability of the proposed WGAIN-GP framework. Overall, the WGAIN-GP model demonstrates accurate imputation performance for the seven missing sensors.

To provide a quantitative assessment, four performance metrics are employed to evaluate the recovery performance of the proposed method. The imputation performance of the WGAIN-GP models in different missing scenarios is summarized in Table 1. The results obtained from these performance metrics demonstrate accurate imputation with low MAE and RMSE values, along with high R2 and accuracy. In conclusion, the proposed WGAIN-GP framework not only achieves accurate imputation but also demonstrates strong generative ability, as confirmed by both qualitative and quantitative assessments.

5.2. The evaluation of modal analysis

Modal identification plays a crucial role in monitoring structural condition and identifying potential damage. However, missing data significantly affects mode recognition and the overall structural state, highlighting the importance of an accurate and stable missing data imputation method. In this study, modal identification was performed on both the original data and the imputed data using the Frequency Domain Decomposition (FDD) technique. FDD decomposes the signal into sinusoidal components with different frequencies and amplitudes, thereby extracting underlying dynamic information. For modal parameter identification, the original and recovered results of seven sensors (J2, J4, J6, J9, J11, J13, and J15) were compared to assess the robustness and accuracy of the proposed method.

To qualitatively and quantitatively analyze the modal identification based on the recovered data, the FDD method was employed. As the bridge structure is a low-frequency system, Fig. 8 demonstrates the accurate identification of six natural frequencies based on the recovered data from the seven sensors. Considering engineering applications, the six modes were inferred by selecting the peak of the singular value. The proposed imputation method exhibits superior performance in terms of natural frequency identification, with an extremely small error rate except for mode 1 (Fig. 8 and Table 2). The identified damping ratios, as shown in Table 2, also exhibit low errors when comparing the original and imputed data. Additionally, mode shapes are effectively identified by comparing the recovered data with the original data, as depicted in Fig. 9. These results highlight the high quality and efficiency of the proposed WGAIN-GP in recovering missing data. To quantify the mode shape comparison, the Modal Assurance Criterion (MAC) is employed as a measure of correlation between two vibration shapes. Table 2 indicates that the MAC values are consistently above 0.995, and even reach 1, demonstrating the accuracy of the proposed method. Overall, the recovered data effectively and accurately identify the modal parameters of the six modes, affirming the superior capacity of the proposed WGAIN-GP model in assessing structural conditions.

$$MAC = \frac{\left|\{\psi_R\}^H\{\psi_I\}\right|^2}{\{\psi_R\}^H\{\psi_R\}\{\psi_I\}^H\{\psi_I\}} \tag{16}$$

where ψ_R and ψ_I are real values representing the mode shape values of the kth mode estimated from the true and reconstructed responses.

5.3. The robustness performance of the proposed WGAIN-GP

To demonstrate the superiority of the proposed WGAIN-GP model, a comparison is made with the previous GAIN and WGAIN-CP models. To ensure a fair comparison, uniform standards are followed in selecting initial hyperparameters and sensor information, which are shown in section 3.3. Fig. 10 clearly illustrates the issues of model collapse and gradient vanishing present in the previous GAIN and WGAIN-CP models, where constant values persist throughout the iteration process. In contrast, the proposed WGAIN-GP model generates available data at the beginning of the iteration and shows a limitation on the loss in different iterations. This highlights the contribution of the Wasserstein loss function and gradient penalty in enhancing the stability and quality of the proposed WGAIN-GP model.

To better illustrate the robustness of the proposed WGAIN-GP model, its performance is compared with WGAIN-CP and GAIN models under different scenarios: random, continuous, and hybrid (a combination of random and continuous) missing data.

Fig. 11 presents the recovery accuracy of the three imputation methods in the presence of random missing data (Type I) occurring in SHM systems. In the case of 50% random missing data, the proposed WGAIN-GP model demonstrates superior recovery results for both high-frequency and low-frequency signals. In contrast, the predicted results of

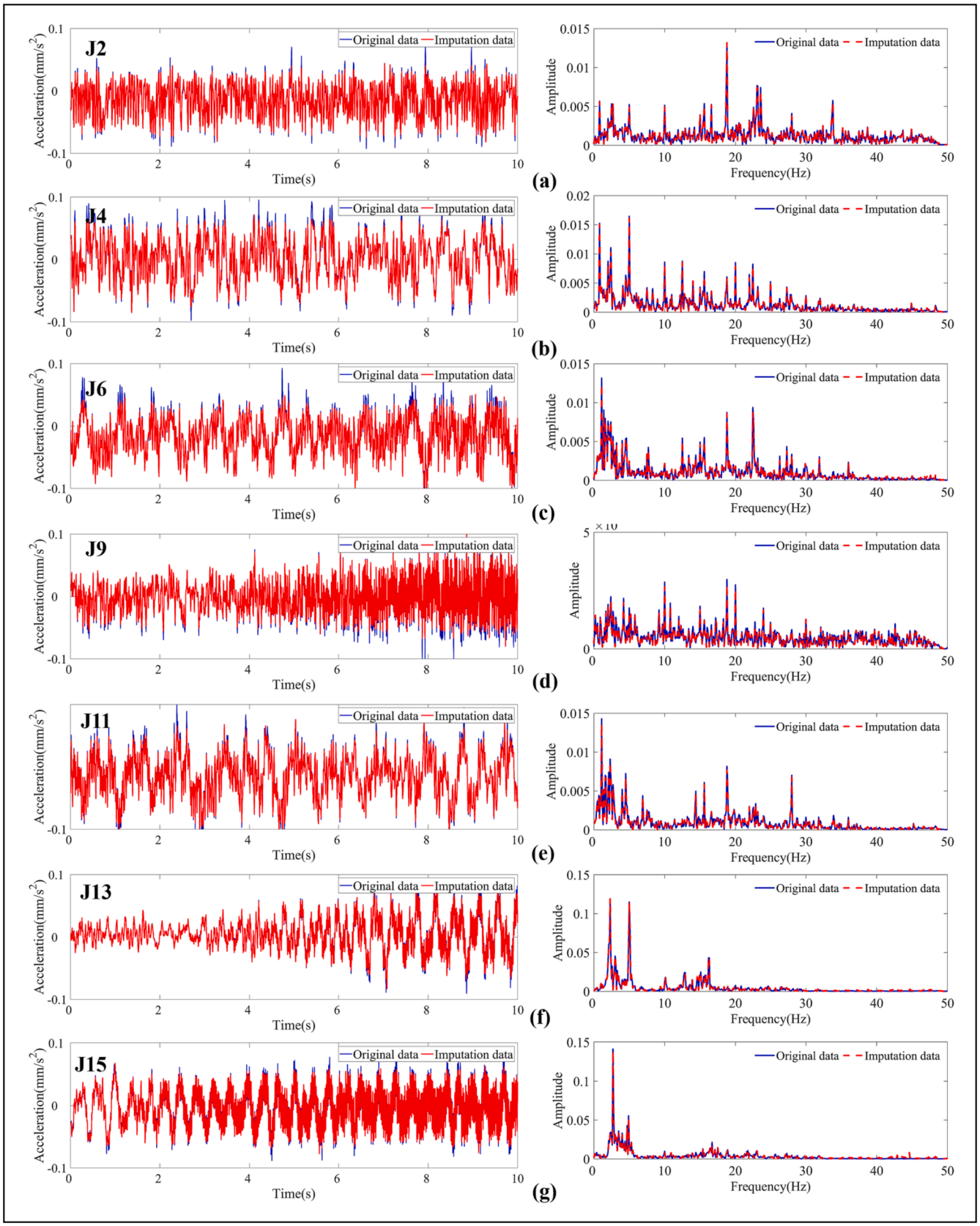

Fig. 5. The imputation results of seven sensors under random missing scenarios: missing rate = 90%.

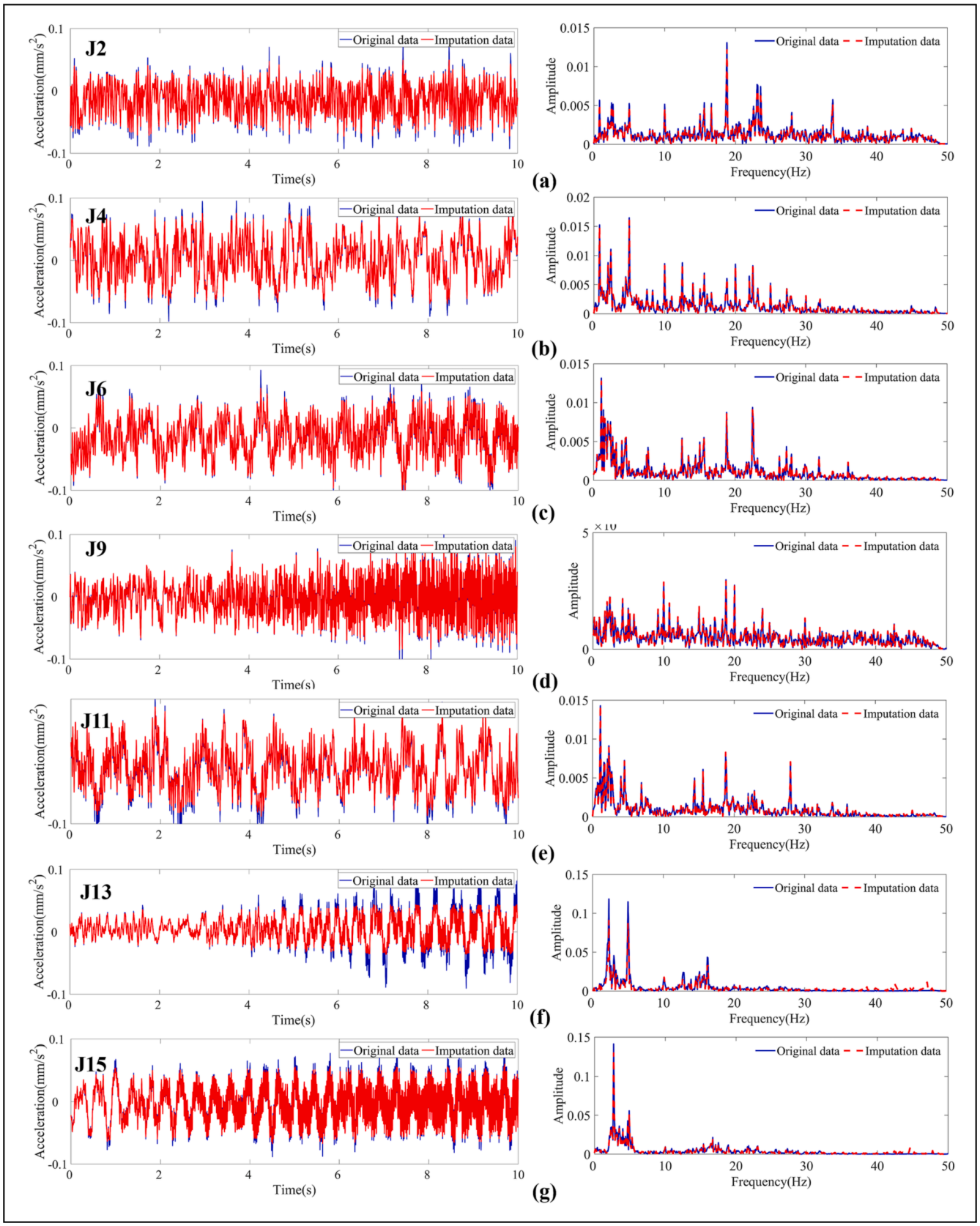

Fig. 6. The imputation results of seven sensors under continuous missing scenarios: missing rate = 90%.

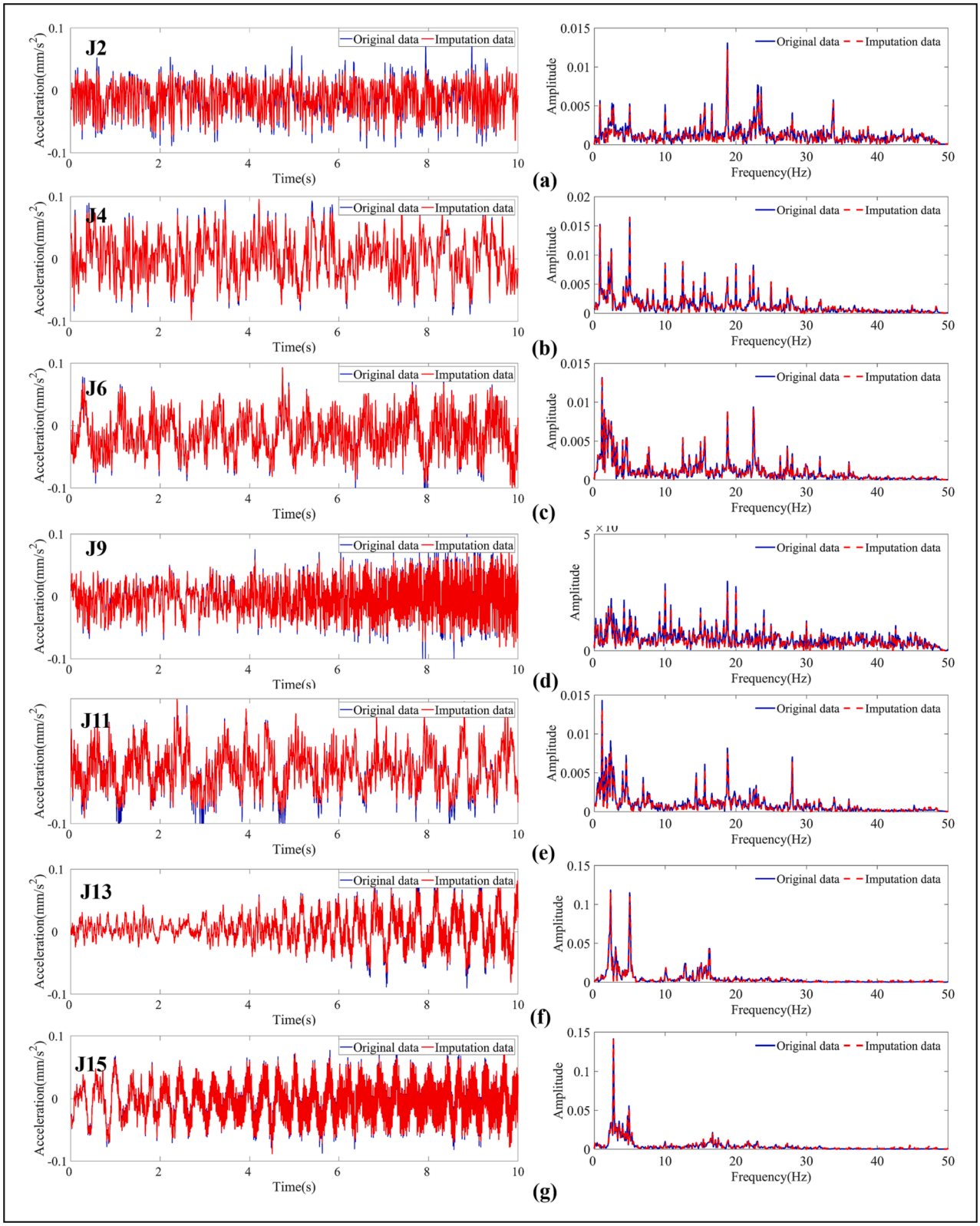

Fig. 7. The imputation results of seven sensors under hybrid missing scenarios (continuous and random): missing rate = 90%.

Table 1
The performance metrics of different missing scenarios with 90% missing ratio.

Missing type	Sensors	MAE	RMSE	R2	Accuracy
Type I	J2	0.021	0.036	0.9845	0.9996
	J4	0.029	0.049	0.9809	0.9986
	J6	0.029	0.053	0.9751	0.9973
	J9	0.01	0.021	0.9866	0.9996
	J11	0.015	0.031	0.9896	0.9856
	J13	0.13	0.224	0.9915	0.9469
	J15	0.172	0.255	0.9897	0.9375
Type II	J2	0.017	0.034	0.9857	0.9996
	J4	0.023	0.037	0.9894	0.999
	J6	0.017	0.029	0.9926	0.9986
	J9	0.075	0.011	0.9962	0.9998
	J11	0.029	0.043	0.9792	0.9797
	J13	0.341	0.785	0.8944	0.9136
	J15	0.179	0.285	0.9872	0.9303
Type III	J2	0.047	0.066	0.9483	0.9993
	J4	0.022	0.029	0.9932	0.9992
	J6	0.018	0.28	0.9928	0.9986
	J9	0.015	0.22	0.9847	0.9995
	J11	0.024	0.044	0.9779	0.979
	J13	0.128	0.234	0.9906	0.9444
	J15	0.133	0.181	0.9949	0.9558

the WGAIN-CP and GAIN models do not align well with the original signal, particularly in high-frequency signals.

When dealing with continuous missing data caused by lost signal contact in SHM systems (referred to as Type II), the recovery accuracy of the three imputation methods is depicted in Fig. 12. Continuous missing data differs significantly from random missing data in terms of imputation implementation, posing a challenge for the proposed WGAIN-GP model. This challenge lies in its ability to not only utilize its own non-missing data but also capture information from other sensors. However, as observed in Fig. 6, the proposed WGAIN-GP model exhibits stable imputation capability when addressing continuous data missing issues with signals of varying frequencies.

When dealing with hybrid missing data caused by a combination of continuous and random missing occurrences in SHM systems (referred to as Type III), the recovery accuracy of the three imputation methods is illustrated in Fig. 13. It is evident that the recovery of Type III data presents a challenging task, as the imputation method needs to effectively incorporate different types of information. However, it is observed that the predicted frequency spikes are internally consistent with the original spikes, indicating the capability of the proposed imputation methods to capture the underlying patterns in the missing data.

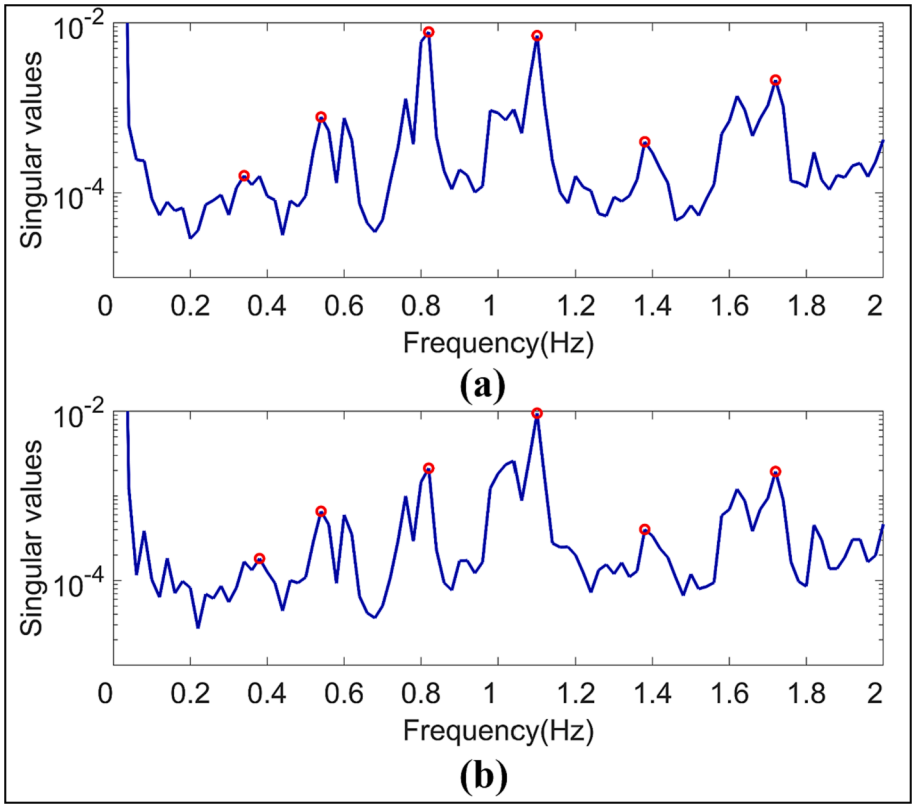

Fig. 8. Modal Frequency Identification results using (a) Measured data, (b)Imputation data.

Table 2
Comparison of the modal identification results under the recovered data of seven sensors.

Mode	Frequency			Damping ratio			MAC
	Real	Imputation	Error	Real	Imputation	Error	
1	0.34	0.38	11.76%	0.016	0.026	62.5%	0.998
2	0.54	0.54	0%	0.021	0.022	4.76%	0.996
3	0.82	0.82	0%	0.01	0.012	20%	0.995
4	1.10	1.10	0%	0.008	0.009	12.5%	0.997
5	1.38	1.38	0%	0.121	0.176	45.45%	1
6	1.72	1.72	0%	0.017	0.018	5.88%	1

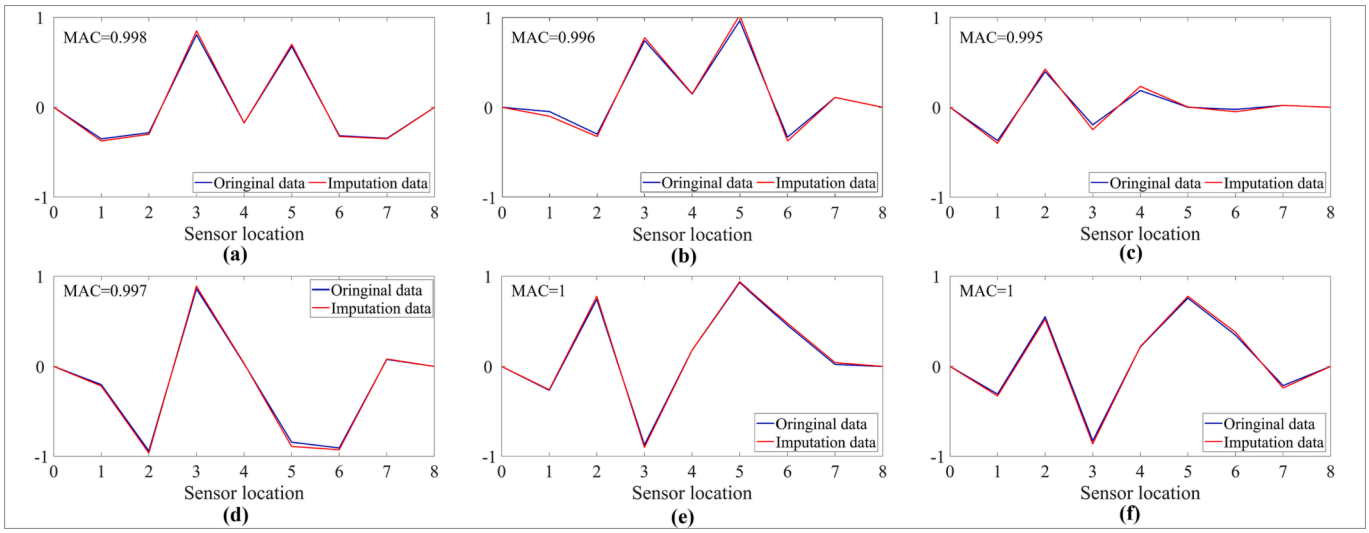

Fig. 9. Modal Shape results using (a) 1st mode shape, (b) 2nd mode shape, (c) 3rd mode shape, (d) 4th mode shape, (e) 5th mode shape, (f) 6th mode shape.

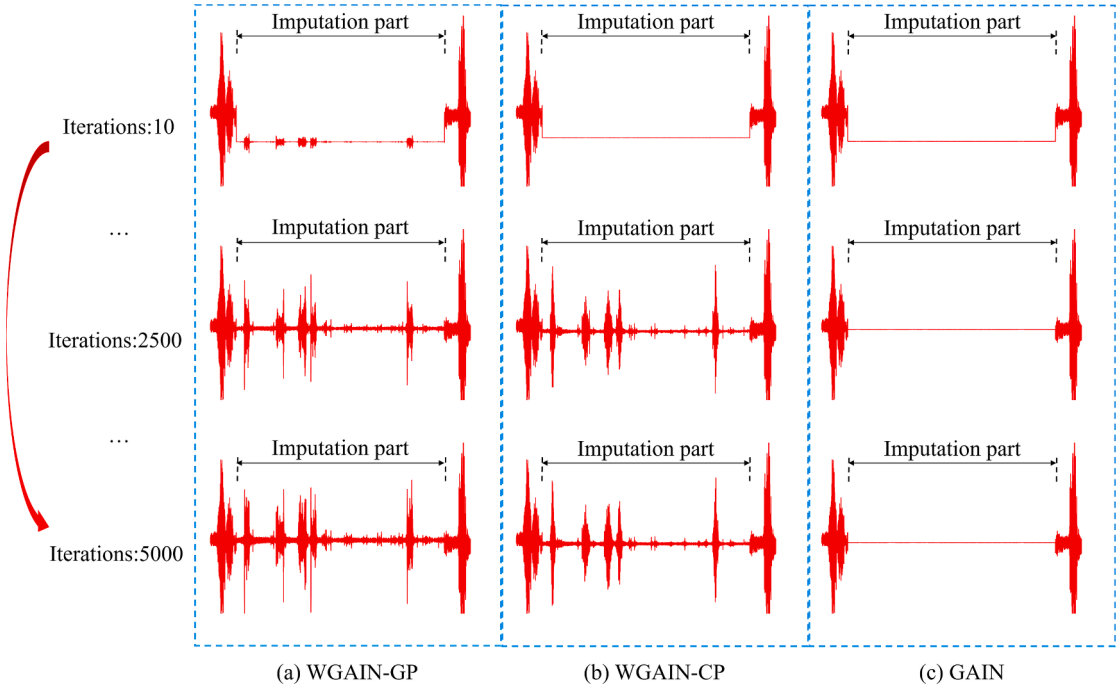

Fig. 10. Train Process of (a) WGAIN-GP (b) WGAIN-CP (c)GAIN.

To comprehensively demonstrate the superiority of the WGAIN-GP framework, quantitative analysis is performed comparing the three imputation models, as shown in Table 3. The WGAIN-GP framework consistently achieves more accurate results and a more stable training process, as indicated by low MAE, RMSE, and high R2 accuracy values. These findings highlight the reliability and robustness of the proposed method as a missing data imputation framework.

6. Conclusion

This paper presents a robust missing data imputation framework, WGAIN-GP, based on a generative adversarial network with Wasserstein distance and gradient penalty. The proposed method is extensively evaluated using measured acceleration data from a long-span highway-railway dual-purpose bridge, demonstrating its effectiveness and robustness in recovering missing data. The implementation results highlight the framework's superior performance across various missing data scenarios, even with a missing data rate of up to 90%. The method's generality is also demonstrated by successfully handling different missing sensors and enabling modal analysis of the bridge's structural state with data recovery.

The assessment of the WGAIN-GP model in different missing scenarios confirms its ability to leverage spatiotemporal correlations among sensors, resulting in accurate imputation. The performance metrics consistently show low mean absolute error (MAE), root mean square error (RMSE), and high R2 accuracy, indicating the reliability and precision of the proposed framework. Furthermore, the WGAIN-GP framework exhibits strong generative ability, as confirmed by both qualitative and quantitative assessments.

Modal identification, a critical aspect of monitoring structural conditions and detecting possible damage, is effectively performed using the proposed imputation method. The recovery results demonstrate

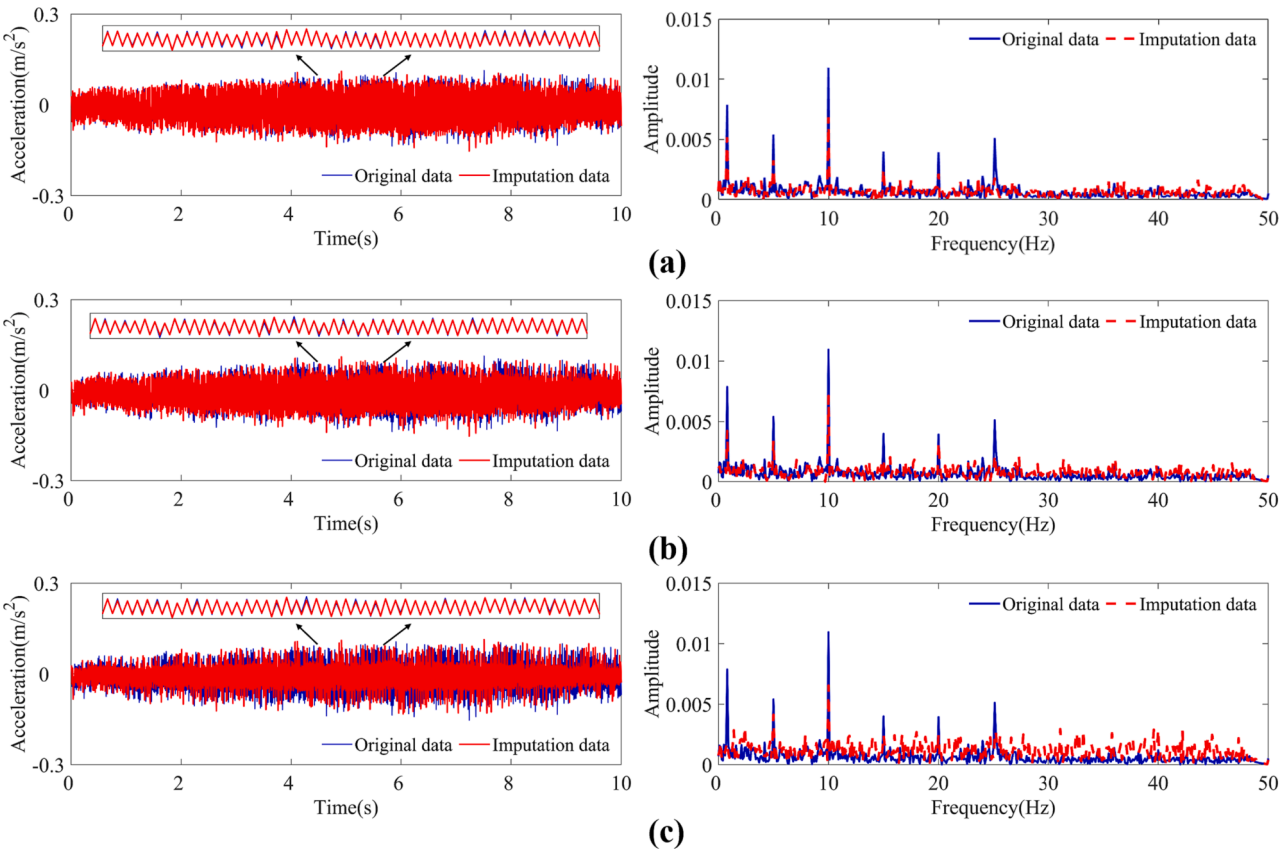

Fig. 11. The imputation results of 50% random missing in time and frequency domain (a) WGAIN-GP (b) WGAIN-CP (c)GAIN.

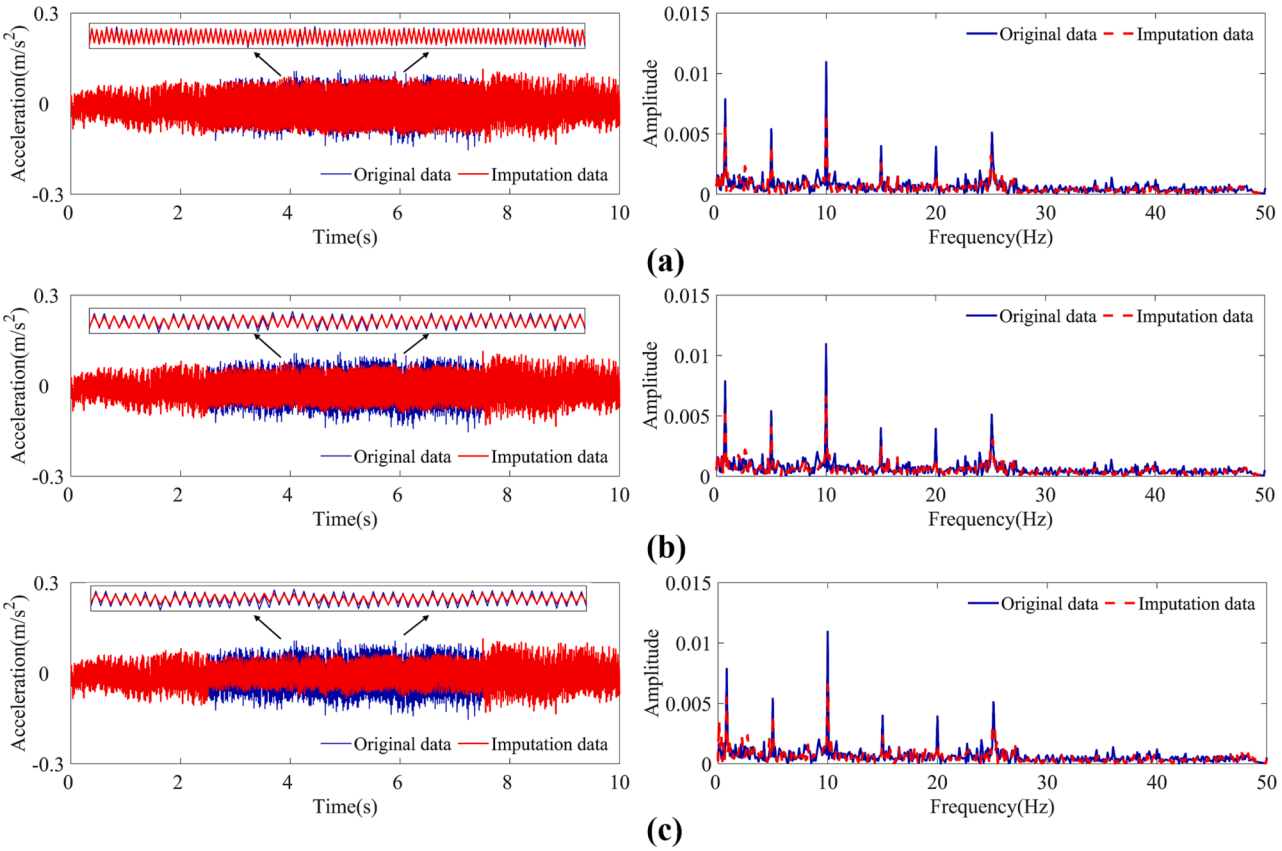

Fig. 12. The imputation results of 50% continuous missing in time and frequency domain (a) WGAIN-GP (b) WGAIN-CP (c)GAIN.

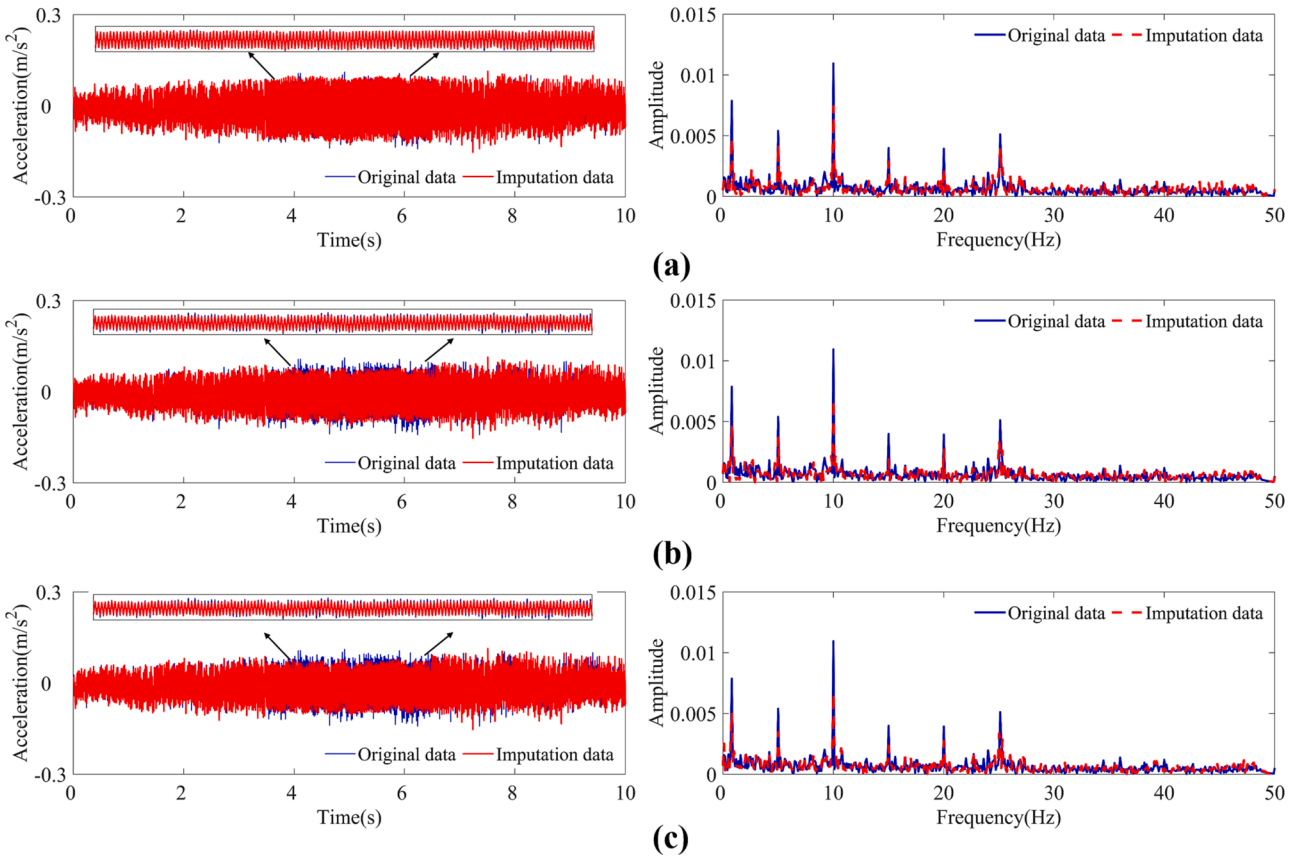

Fig. 13. The imputation results of 30% continuous and 20% random missing in time and frequency domain (a) WGAIN-GP (b) WGAIN-CP (c)GAIN.

Table 3
The performance metrics of different missing imputation methods with 50% missing ratio.

Missing type	Model	MAE	RMSE	R2	Accuracy
Type I	WGAIN-GP	0.052	0.099	0.9828	0.998
	WGAIN-CP	0.093	0.101	0.9723	0.997
	GAIN	0.184	0.289	0.9546	0.956
Type II	WGAIN-GP	0.059	0.108	0.9798	0.996
	WGAIN-CP	0.159	0.138	0.9697	0.973
	GAIN	0.16	0.184	0.9596	0.955
Type III	WGAIN-GP	0.062	0.111	0.9787	0.997
	WGAIN-CP	0.094	0.125	0.9671	0.975
	GAIN	0.132	0.193	0.9542	0.961

accurate identification of natural frequencies, with minimal error except for mode 1. Additionally, the damping ratios exhibit low error between original and imputed data. The mode shapes are successfully identified, further affirming the high quality and efficiency of the proposed WGAIN-GP framework. The modal assurance criterion (MAC) values consistently exceed 0.995, indicating excellent accuracy in comparing mode shapes between the recovered and original data. Overall, the proposed framework excels in assessing the structural conditions by effectively and accurately identifying the modal parameters.

A comparison with other imputation models, GAIN and WGAIN-CP, reveals the limitations of those models, such as model collapse and gradient vanishing. In contrast, the proposed WGAIN-GP model generates reasonable data from the start of iterations, demonstrating the stability and quality achieved through the Wasserstein loss function and gradient penalty. Quantitative analysis confirms the superior performance of the WGAIN-GP framework, with lower MAE and RMSE, higher R2 accuracy, and a more stable training process. These results affirm the reliability and robustness of the proposed method as a missing data imputation framework.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by the National Key R&D Program of China (2021YFE0112200), the Japan Society for Promotion of Science (Kakenhi No. 18K04438), the Tohoku Institute of Technology research Grant, and the Science and Technology Research Project of Education Department of Jiangxi Province (Grant No. GJJ2200682).

References

[1] Worden K, Manson G. The application of machine learning to structural health monitoring. Philos Trans R Soc A Math Phys Eng Sci 2007;365(1851):515–37. https://doi.org/10.1098/rsta.2006.1938.

[2] Sohn H. Effects of environmental and operational variability on structural health monitoring. Philos Trans R Soc A Math Phys Eng Sci 2007;365(1851):539–60. https://doi.org/10.1098/rsta.2006.1935.

[3] Dong W, Fong DYT, Yoon J-S, Wan EYF, Bedford LE, Tang EHM, et al. Generative adversarial networks for imputing missing data for big data clinical research. BMC Med Res Method 2021;21(1). https://doi.org/10.1186/s12874-021-01272-3.

[4] Ou J, Li H. Structural health monitoring in mainland china: Review and future trends. Struct Heal Monit 2010;9(3):219–31. https://doi.org/10.1177/1475921710365269.

[5] B. Chapuis, "Introduction to Structural Health Monitoring," pp. 1–11, 2018, doi: 10.1007/978-3-319-69233-3_1.

[6] Cross EJ, Koo KY, Brownjohn JMW, Worden K. Long-term monitoring and data analysis of the Tamar Bridge. Mech Syst Sig Process 2013;35(1–2):16–34. https://doi.org/10.1016/j.ymssp.2012.08.026.

[7] Stekhoven DJ, Bühlmann P. Missforest-Non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28(1):112–8. https://doi.org/10.1093/bioinformatics/btr597.

[8] Oh BK, Glisic B, Kim Y, Park HS. Convolutional neural network–based data recovery method for structural health monitoring. Struct Heal Monit 2020;19(6):1821–38. https://doi.org/10.1177/1475921719897571.

[9] Du J, Chen H, Zhang W. A deep learning method for data recovery in sensor networks using effective spatio-temporal correlation data. Sens Rev 2019;39(2):208–17. https://doi.org/10.1108/SR-02-2018-0039.

[10] Li Y, Bao T, Chen H, Zhang K, Shu X, Chen Z, et al. A large-scale sensor missing data imputation framework for dams using deep learning and transfer learning strategy. Meas J Int Meas Confed 2021;178:109377.

[11] García-Laencina PJ, Sancho-Gómez JL, Figueiras-Vidal AR. Pattern classification with missing data: A review. Neural Comput & Applic 2010;19(2):263–82. https://doi.org/10.1007/s00521-009-0295-6.

[12] Goulet J-A, Koo K. Empirical Validation of Bayesian Dynamic Linear Models in the Context of Structural Health Monitoring. J Bridg Eng 2018;23(2):05017017. https://doi.org/10.1061/(asce)be.1943-5592.0001190.

[13] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1–67. https://doi.org/10.18637/jss.v045.i03.

[14] Zhang N, Ding S, Zhang J, Xue Y. An overview on Restricted Boltzmann Machines. Neurocomputing 2018;275:1186–99. https://doi.org/10.1016/j.neucom.2017.09.065.

[15] Lin WC, Tsai CF, Zhong JR. Deep learning for missing value imputation of continuous data and the effect of data discretization. Knowledge-Based Syst 2022;239:108079. https://doi.org/10.1016/j.knosys.2021.108079.

[16] P. Clavier, "Sum-Product Network in the context of missing data," 2020.

[17] Arul M, Kareem A, Kwon DK. Identification of vortex-induced vibration of tall building pinnacle using cluster analysis for fatigue evaluation: application to Burj Khalifa. J Struct Eng 2020;146(11):04020234. https://doi.org/10.1061/(asce)st.1943-541x.0002799.

[18] Hawthorne G, Elliott P. Imputing cross-sectional missing data: Comparison of common techniques. Aust N Z J Psychiatry 2005;39(7):583–90. https://doi.org/10.1111/j.1440-1614.2005.01630.x.

[19] Richman MB, Trafalis TB, Adrianto I. Missing data imputation through machine learning algorithms. Artif Intell Methods Environ Sci 2009:153–69. https://doi.org/10.1007/978-1-4020-9119-3_7.

[20] Ni YQ, Li M. Wind pressure data reconstruction using neural network techniques: A comparison between BPNN and GRNN. Meas J Int Meas Confed 2016;88:468–76. https://doi.org/10.1016/j.measurement.2016.04.049.

[21] Li Y, Bao T, Chen Z, Gao Z, Shu X, Zhang K. A missing sensor measurement data reconstruction framework powered by multi-task Gaussian process regression for dam structural health monitoring systems. Measurement 2021;186:110085.

[22] Chen Z, Bao Y, Li H, Spencer BF. LQD-RKHS-based distribution-to-distribution regression methodology for restoring the probability distributions of missing SHM data. Mech Syst Sig Process 2019;121:655–74. https://doi.org/10.1016/j.ymssp.2018.11.052.

[23] Ren Pu, Chen X, Sun L, Sun H. Incremental Bayesian matrix/tensor learning for structural monitoring data imputation and response forecasting. Mech Syst Sig Process 2021;158:107734.

[24] Wan H-P, Ni Y-Q. Bayesian modeling approach for forecast of structural stress response using structural health monitoring data. J Struct Eng 2018;144(9):1–12. https://doi.org/10.1061/(asce)st.1943-541x.0002085.

[25] Wan HP, Sen Dong G, Luo Y, Ni YQ. An improved complex multi-task Bayesian compressive sensing approach for compression and reconstruction of SHM data. Mech Syst Sig Process 2022;167(PA):108531. https://doi.org/10.1016/j.ymssp.2021.108531.

[26] Wan HP, Ni YQ. Bayesian multi-task learning methodology for reconstruction of structural health monitoring data. Struct Heal Monit 2019;18(4):1282–309. https://doi.org/10.1177/1475921718794953.

[27] Li L, Zhou H, Liu H, Zhang C, Liu J. A hybrid method coupling empirical mode decomposition and a long short-term memory network to predict missing measured signal data of SHM systems. Struct Heal Monit 2021;20(4):1778–93. https://doi.org/10.1177/1475921720932813.

[28] Xia Y, Lei X, Wang P, Sun L. A data-driven approach for regional bridge condition assessment using inspection reports. Struct Control Heal Monit 2022;29(4):1–18. https://doi.org/10.1002/stc.2915.

[29] Fan G, Li J, Hao H, Xin Yu. Data driven structural dynamic response reconstruction using segment based generative adversarial networks. Eng Struct 2021;234:111970.

[30] Fan G, He Z, Li J. Structural dynamic response reconstruction using self-attention enhanced generative adversarial networks. Eng Struct 2023;276:115334. https://doi.org/10.1016/j.engstruct.2022.115334.

[31] Jiang H, Wan C, Yang K, Ding Y, Xue S. Continuous missing data imputation with incomplete dataset by generative adversarial networks–based unsupervised learning for long-term bridge health monitoring. Struct Heal Monit 2022;21(3):1093–109.

[32] Hou J, Jiang H, Wan C, Yi L, Gao S, Ding Y, et al. Deep learning and data augmentation based data imputation for structural health monitoring system in multi-sensor damaged state. Meas J Int Meas Confed 2022;196. https://doi.org/10.1016/j.measurement.2022.111206. 111206.

[33] Jiang K, Han Q, Du X. Lost data neural semantic recovery framework for structural health monitoring based on deep learning. Computer Aided Civil Eng 2022;37(9):1160–87.

[34] Chen C, Tang L, Lu Y, Zhou L, Liu Z, Liu Y, et al. Temperature-induced response reconstruction method based on DL-AR model and attention mechanism. Structures 2023;50:359–72.

[35] Gui J, Sun Z, Wen Y, Tao D, Ye J. A review on generative adversarial networks: algorithms, theory, and applications. IEEE Trans Knowl Data Eng 2021;14(8):1. https://doi.org/10.1109/tkde.2021.3130191.

[36] L. Weng, "From GAN to WGAN," 2019, [Online]. Available: http://arxiv.org/abs/1904.08994.

[37] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM 2020;63(11):139–44.

[38] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, [Online]. Available: http://arxiv.org/abs/1701.07875.

[39] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A. Improved training of wasserstein GANs. Adv Neural Inf Process Syst 2017;vol. 2017-Decem:5768–78.

[40] Gao S, Zhao W, Wan C, Jiang H, Ding Y, Xue S. Missing data imputation framework for bridge structural health monitoring based on slim generative adversarial networks. Meas J Int Meas Confed 2022;204(October):112095. https://doi.org/10.1016/j.measurement.2022.112095.

[41] Sen Dong G, Wan HP, Luo Y, Todd MD. A fast sparsity-free compressive sensing approach for vibration data reconstruction using deep convolutional GAN. Mech Syst Signal Process 2023;188(November 2022):109937. https://doi.org/10.1016/j.ymssp.2022.109937.

[42] L. Mescheder, A. Geiger, and S. Nowozin, "Which training methods for GANs do actually converge?," *35th Int. Conf. Mach. Learn. ICML 2018*, vol. 8, pp. 5589–5626, 2018.

[43] D. T. Neves, M. G. Naik, and A. Proen, "Novel GAN Methods for Missing Data Imputation," pp. 1–14, 2021, doi: 10.1007/978-3-030-77961-0.

[44] J. Yoon, J. Jordon, and M. Van Der Schaar, "GAIN: Missing data imputation using generative adversarial nets," *35th Int. Conf. Mach. Learn. ICML 2018*, vol. 13, pp. 9042–9051, 2018.