scientific reports

Check for updates

OPEN

A novel damage detection method based on sequential iteration and Gaussian mixture model for structural health monitoring under environmental effects

Jie-zhong Huang[1,2,4], Jian Yang[1], Dong-sheng Li[1,3✉], Wei-chen Kong[5] & Ya-fei Wang[6]

Environmental effects often cause variability in dynamic features, obscuring actual damage indicators and leading to false alarms in damage detection. The Gaussian mixture model (GMM) based method is an effective solution, but challenges such as selecting initial model parameters and determining the optimal number of Gaussian components can hinder its performance. To address these challenges, we propose a two-step method that combines sequential iteration with the GMM approach. In the first step, sequential iteration is employed to determine initial model parameters and the optimal number of Gaussian components for a reliable GMM. In the second step, the expectation-maximization (EM) algorithm is used to establish the GMM, clustering the training data into local subsets. For each subset, the Mahalanobis squared distance (MSD) between each sample point and the center of its Gaussian component is calculated. This distance is used to create a novelty index based on the minimum Mahalanobis squared distance (MMSD), facilitating effective damage detection by the statistical control chart. Moreover, generalized extreme value distribution modeling method is presented to determine an accurate control limit. We validate our method using real data from two bridges, demonstrating its effectiveness through comparative analysis.

Keywords Structural health monitoring (SHM), Environmental and operational conditions, Damage detection, Gaussian mixture model, Z24 Bridge

Bridge structural health monitoring (SHM) plays a vital role in ensuring the safety and longevity of transportation infrastructure. Over time, bridges are subjected to various factors such as traffic loads, environmental conditions, and natural aging, which can cause deterioration. Early detection of damage through SHM enables timely maintenance, reducing the risk of catastrophic failures. By continuously assessing the condition of bridge components, the need for costly repairs can be minimized, and service life can be extended[1,2]. This proactive approach ultimately enhances public safety, reduces disruptions to transportation, and promotes sustainable infrastructure management.

In recent years, there has been extensive research on vibration-based damage identification methods within the field of SHM[3]. These methods operate on the principle that damage can induce alterations in the dynamic characteristics of structures, including frequency and mode shapes. Subsequently, according to changes in the identified dynamic features, structural damage can be detected through the application of inverse problem-solving techniques[4]. However, in practical engineering applications, the dynamic characteristics of structures are not solely influenced by damage; they are also significantly impacted by diverse environmental and operational conditions, such as temperature[5,6] humidity[7] wind[8] and traffic loads[9]. These influences can even completely

[1]Department of Civil and Intelligent Construction Engineering, Shantou University, Shantou, Guangdong Province, China. [2]Anhui Provincial International Joint Research Center of Data Diagnosis and Smart Maintenance on Bridge Structures, Chuzhou, Anhui Province, China. [3]Shantou Key Laboratory of Offshore Wind Energy, Shantou, Guangdong Province, China. [4]State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian, Liaoning Province, China. [5]Civil and Environment Engineering School, University of Science and Technology Beijing, Beijing, China. [6]Key Laboratory for Health and Safety of Bridge Structures, Wuhan, Hubei Province, China. ✉email: lids@stu.edu.cn

mask the changes in dynamic characteristics caused by damage, thereby making accurate detection of damage occurrence challenging if environmental factors are not duly considered[10].

Accordingly, researchers have proposed various methods to effectively remove the environmental effects from the damage identification process. Depending on whether the environmental parameters are measured, these methods can be categorized as input-output and output-only methods[11]. Input-output methods attempt to establish an explicit relationship between environmental variables and dynamic characteristics to quantify the impact of environmental variables on dynamic characteristics. Common methods include linear regression[12,13] polynomial regression[14] bilinear regression[15] support vector regression[16] random forest models[17] Gaussian process regression[18] and artificial neural networks[19] among others. These methods use environmental variables and the monitored dynamic characteristics as the model's inputs and outputs, respectively, making it easier to interpret the influence of environmental variables on dynamic characteristics. However, their drawback lies in the requirement for comprehensive measurement of environmental variables, which is often difficult to obtain in practice[20]. Furthermore, once the input-output model is established, the sensors must remain in their original positions, as any failure in a sensor can lead to false-positive or false-negative outcomes[21].

Compared to input-output methods, output-only methods do not directly measure environmental information; instead, they rely solely on the measured dynamic characteristics to learn the implicit relationship between environmental variables and dynamic characteristics[22]. The fundamental concept behind output-only methods is to transform the damaged features into a feature space through a suitable transformation[23]. This space yields new features that are insensitive to environmental operating variables but retain sensitivity to damage. Over recent years, output-only methods have gained widespread attention[24–26] and several commonly employed methods in this context include PCA and its variants[27–33] cointegration and its variants[34–36] factor analysis[37] GMM[38,39] Mahalanobis squared-distance[37] singular spectrum analysis[24] and auto-associative neural networks[40,41] nonlinear narrow dimension techniques[42]. More recently, Mousavi et al. developed a nonlinear method based on variational mode decomposition, cointegration, and recurrent neural network. The superiority of the proposed method was demonstrated by an experimental example of the Z24 bridge[43]. Wah et al.[21] developed a multiple-regression-based damage detection method under changing environmental and operational conditions. This method employs natural frequencies as both the independent and dependent variables and utilizes the difference in fits to identify influential observations and remove outlier measurements. Peng et al.[44] proposed a novel structural damage detection based on manifold learning, utilizing the phase space representation of vibration characteristics to detect structural damage under varying environmental and operational conditions. Sarmadi et al.[45] introduced a probabilistic data self-clustering method based on semi-parametric extreme value theory for damage detection. This method effectively handles environmental and operational variability while providing an integrated framework for determining the damage index and decision threshold.

Moreover, some machine learning or deep learning methods are used to address the temperature compensation problem in SHM. Sawant et al.[46] proposed an unsupervised, temperature-compensated damage localization method for guided wave-based structural health monitoring (GW-SHM) using transfer learning from a convolutional autoencoder, achieving improved accuracy and robustness to environmental variations compared to supervised approaches, with fewer trainable parameters and without the need for pre-processed signals or material property knowledge. Du et al.[47] developed an attention-based multi-task network with a modified U-Net architecture for accurate bolt loosening detection and temperature compensation in multi-bolt connections, validated experimentally on a simulated aircraft structure. Kashyap et al.[48] presented a TinyML-based unsupervised learning framework for GW-SHM using a lightweight neural network, enabling real-time damage identification and localization in a honeycomb composite sandwich structure under varying temperature conditions. Sawant et al.[49] proposed a CNN-based supervised framework for automated feature extraction, coupled with a Gaussian mixture model for temperature compensation and damage localization in guided wave-based structural health monitoring.

Gaussian Mixture Model is a highly effective probabilistic statistical tool for characterizing data distributions. The GMM-based methods have gained significant traction for detecting structural damage under changing environmental and operational conditions. For example, Figueiredo et al.[38,39] utilized GMM to model data distribution under varying normal conditions, employing the smallest MSD from each component for damage detection. This approach was successfully demonstrated using daily standard data sets from the Z24 Bridge. Other researchers have also explored the use of GMM in structural health monitoring. Kullaa et al.[50] and Zang et al.[51] applied GMM to cluster modal frequencies into several local subdomains. They then used minimum mean square error estimation and principal component analysis (PCA) to mitigate the effects of environmental and operational variations (EOV) on the subdomain data. Zhang et al.[52] introduced a GMM-based method to quantify fatigue crack size, establishing a comprehensive baseline GMM database under varying temperatures and detecting cracks through the disparity between the baseline and test GMMs. Ren et al.[53] combined GMM with a delay-and-sum method to achieve reliable damage monitoring and imaging localization for structures under variable conditions. More recently, Daneshvar et al.[54] proposed an unsupervised damage detection method that employs GMM to provide local information for a deep reinforcement learning algorithm. This method reconstructs local training subsets using subdictionaries and uses an anomaly detector based on reconstruction error for damage detection. Beyond bridge structures[55,56] GMM has been applied to damage and anomaly detection in various other domains, including gas turbines[57] aircraft wings[58] and wind turbines[59].

Although GMM-based damage detection methods have proven effective, two important issues in mixture modeling need to be addressed: the selection of initial model parameters and the determination of the number of Gaussian components. (1) The first issue concerns the selection of initial model parameters. The common method to fit a GMM is through maximum likelihood estimation using the EM algorithm. However, the EM algorithm is a local optimization method and is highly sensitive to initialization. Due to the non-unimodal nature of the likelihood function in a mixture model, the EM algorithm may struggle to reach the optimal solution if the

initial parameters deviate significantly from the actual model parameters. (2) The second issue is selecting the number of Gaussian components. The Bayesian Information Criterion (BIC) is commonly used to estimate this number[60]. The optimal number of components is indicated by the lowest BIC value, balancing a good fit to the data with the avoidance of overfitting. However, the EM algorithm's sensitivity to initial model parameters can cause variations in BIC values for the same number of components when different initial values are chosen. This instability makes determining the appropriate number of Gaussian components more complex[61,62].

Figueiredo et al.[63] also highlight these two challenges in GMMs: sensitive to initialization and selection of the number of Gaussian components. This paper proposed an unsupervised algorithm for learning finite mixture models from multivariate data, which integrates model estimation and selection into a single process, avoiding the issue of convergence to singular estimates at the boundary of the parameter space. Qiu et al.[58] proposed an Improved Density Peaks Clustering (IDPC)-based EM algorithm for constructing an adaptive GMM. The algorithm improves the initialization process of GMM parameters by adaptively searching for probability density peaks of Guided Wave (GW) Damage Features (DFs), reducing reliance on experience-dependent parameters. Moreover, to ensure the accuracy of the classification results, Wang et al.[64] and Yi et al.[65] used the K-means clustering algorithm to estimate the initial parameters of the GMM.

Although similar methods have been proposed for determining the initial value of GMM, they have certain disadvantages. For example, in the K-means clustering-based GMM method[64,65] to determine the number of clusters for K-means clustering, the BIC method is usually used. However, the variations in BIC values for the same number of components will cause the instability in determining the appropriate number of clusters. In the density peak clustering-based adaptive GMM method[58] the ratio k in the algorithm should be empirically determined, limiting the application of this damage detection method. Moreover, for the unsupervised algorithm for learning finite mixture models[63] it faces the challenge when dealing with mixtures that have components with very different weights, potentially leading to the premature elimination of smaller components or improper overlap of heavier components.

In this paper, we propose a method that integrates a sequential iteration operation into the GMM framework to address two critical issues: (1) the automatic determination of the number of Gaussian components, and (2) the enhancement of the reliability of the initial model parameters for GMM. Our method follows a systematic procedure. First, a sequential iteration approach is employed to analyze the modal frequencies under varying environmental conditions. This approach automatically divides the data into multiple clusters that conform to Gaussian distributions. These clusters provide the mean and covariance of each Gaussian component, which are then used as the initial parameters for the GMM, thereby improving their reliability. Next, the EM algorithm is applied iteratively to refine these parameters until an optimal solution is obtained, ensuring effective optimization of the GMM parameters. Finally, the MSD between the damage feature and each Gaussian component is calculated. Using this information, a statistical process control chart is established for damage detection based on the minimum MSD, and the generalized extreme value distribution modeling method is used to determine an appropriate threshold. To validate the effectiveness of the proposed method, we utilize real data from two bridge structures: a wooden truss bridge and the Z24 Bridge. The damage detection results indicate that the proposed method significantly reduces the occurrence of false alarms, demonstrating superior performance compared to conventional methods in real-world scenarios.

Conventional Gaussian mixture model

GMM is a powerful statistical tool that accurately describes the distribution of real data by employing a weighted linear combination of multiple Gaussian probability density functions[66]. In the context of a dataset $X_{n \times d} = \{x_1, \ldots, x_n\}^{\mathrm{T}}$, where n represents the number of samplings and d denotes the dimension of damage features, GMM assumes that the data can be effectively represented as a mixture of K multivariate Gaussian distributions. Consequently, the probability distribution of X adheres to the principles of the GMM model, the probability density function of X can be derived by summing the weighted Gaussian probability density functions of the K components. This can be mathematically expressed as follows[58]:

$$p(x_i|\Theta) = \sum_{k=1}^{K} \alpha_k \, \mathrm{N}(X|\mu_k, C_k) \tag{1}$$

with parameters

$$\Theta = (\alpha_1, \cdots, \alpha_K; \mu_1, \cdots, \mu_K; C_1, \cdots, C_K) \tag{2}$$

where K is the number of Gaussian components, α_k represents the weight of the k-th Gaussian component which satisfies the constraint that $\sum_{k=1}^{K} \alpha_k = 1$ and $0 \leq \alpha_k \leq 1$. $\mu_k \in R^d$ and $C_k \in R^{d \times d}$ are the mean and covariance matrix of the k-th Gaussian component, respectively. $\mathrm{N}(X|\mu_k, C_k)$ represents the probability density function of the normal distribution:

$$N(x_i \mid \mu_k, C_k) = \frac{1}{(2\pi)^{\frac{d}{2}}|C_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_i - \mu_k)^{\mathrm{T}} C_k^{-1}(x_i - \mu_k)} \tag{3}$$

To estimate the unknown parameters, the maximum likelihood estimation based on the EM algorithm is usually used. The log likelihood function to be maximized is defined as follows,

$$L(X|\Theta) = \log \prod_{i=1}^{n} p(x_i|\Theta) = \sum_{i=1}^{n} \log \sum_{k=1}^{K} \alpha_k \, \mathrm{N}(X|\mu_k, C_k) \tag{4}$$

The EM algorithm is an iterative process that comprises two main steps[58]: (a) the E-step, which involves estimating the posterior probability using Eq. (5); and (b) the M-step, where the Gaussian component model parameters are updated based on the posterior probability, utilizing Eqs. (6)-(8). Initially, the parameters in Θ are randomly selected, and the E-step and M-step are iterated repeatedly until the estimated Θ remains almost unchanged. This convergence indicates that Θ has reached a local optimum value. Finally, the model parameters Θ of the GMM that maximize the log-likelihood function $L(X|\Theta)$ are obtained.

$$\Psi_{ik}(x_i|\Theta) = \frac{\alpha_k \, N(x_i|\mu_k, C_k)}{\sum_{j=1}^{K} \alpha_j \, N(x_i|\mu_j, C_j)} \tag{5}$$

$$\alpha_k = \frac{1}{n} \sum_{i=1}^{n} \Psi_{ik}(x_i|\Theta) \tag{6}$$

$$\mu_k = \frac{\sum_{i=1}^{n}[\Psi_{ik}(x_i|\Theta)x_i]}{\sum_{i=1}^{n} \Psi_{ik}(x_i|\Theta)} \tag{7}$$

$$C_k = \frac{\sum_{i=1}^{n}[\Psi_{ik}(x_i|\Theta)(x_i - \mu_k)(x_i - \mu_k)^T]}{\sum_{i=1}^{n} \Psi_{ik}(x_i|\Theta)} \tag{8}$$

However, the EM algorithm is a local optimization method due to the non-convex likelihood surface and iterative refinement, it may encounter difficulties in achieving the optimal solution when the initial parameters deviate significantly from the actual model parameters[63]. To illustrate the impact of randomly selecting initial parameters on the conventional GMM method, consider a clustering example using a 2-dimensional simulation dataset. When setting the number of model components to 13, the EM algorithm is applied to model the dataset multiple times using different initial values. Figure 1 showcases the modeling results obtained from three separate initial parameters. It is evident that the clustering outcomes obtained by the EM algorithm vary due to the usage of three different sets of initial parameters, with only one correct clustering result. This highlights that the random selection of initial parameters can lead to instability in the GMM model established by the EM algorithm.

In addition to selecting appropriate initial parameters, determining the appropriate number of Gaussian components is also crucial for building a reliable GMM. The BIC is commonly used for this purpose. It involves using the EM algorithm to establish GMMs with varying numbers of components, selecting the number of components corresponding to the minimum BIC value as the optimal choice. However, due to the randomness associated with initial parameter selection, BIC values can fluctuate for the same number of Gaussian components. To illustrate this variability, Fig. 2 shows the changes in BIC values with the number of components for the simulation data. In this illustration, the EM algorithm is executed five times with different initial parameters for each component. It is evident that the calculated BIC value varies significantly for a given number of Gaussian components due to different initial parameters. This instability in the BIC curve results from the EM algorithm's sensitivity to random initialization, which reduces the reliability of the BIC method for determining the optimal number of components.

Improved GMM damage detection method
Improved GMM based on sequential iteration

To address the challenges in conventional GMM, we incorporate a sequential iteration operation to obtain stable and reasonable initial parameters[67]. This sequential iteration method automatically partitions the data into several clusters that adhere to Gaussian distributions, with each cluster serving as an initial Gaussian component of the GMM. The mean and covariance of each Gaussian component, derived from the sequential iteration


(a)　　　　　　　　　(b)　　　　　　　　　(c)
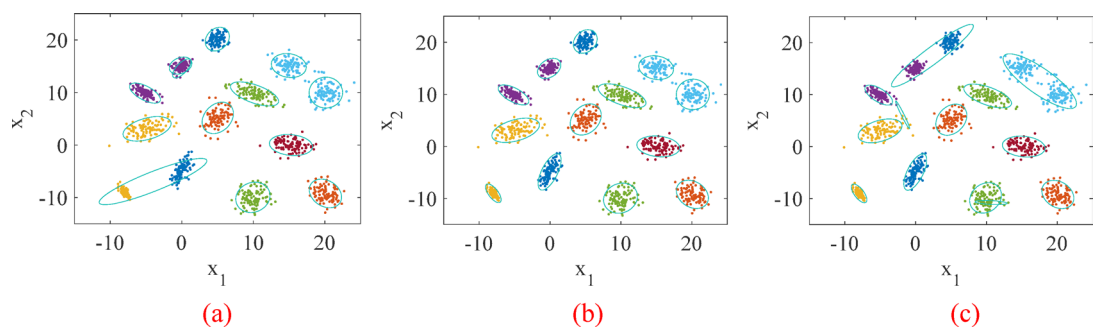
Fig. 1. Three clustering results based on the conventional GMM method by randomly using 3 sets of initial parameters. Table 1 shows the three groups of selected initial sample points. For each set of initial parameters, the initial weight of the k-th Gaussian component is set as $\alpha_k^0 = 1/13$, the initial covariance matrix of the k-th Gaussian component is set as $C_k^0 = \begin{bmatrix} 86.57 & 1.89 \\ 1.89 & 97.09 \end{bmatrix}$, the initial mean of the k-th Gaussian component is listed in Table 1.

k-th Gaussian component	The 1st set		The 2nd set		The 3rd set	
	x_1	x_2	x_1	x_2	x_1	x_2
1	14.99	-1.74	-5.40	3.78	-7.18	11.66
2	-0.46	15.25	-8.26	-8.85	-0.27	14.60
3	-6.74	2.51	14.33	15.15	-8.30	-8.91
4	-5.72	0.74	20.69	8.33	18.05	8.54
5	20.48	10.83	-3.52	2.04	10.95	9.49
6	16.13	13.94	3.99	5.36	-4.59	9.87
7	21.18	9.71	16.07	14.93	-6.23	9.36
8	-1.32	5.32	9.10	10.49	8.47	-8.65
9	4.73	9.00	4.11	18.78	20.12	-8.44
10	9.08	-10.19	16.10	16.32	10.69	8.83
11	4.83	20.59	-8.43	-8.94	-0.33	-6.61
12	-4.45	2.38	-5.16	1.62	15.81	-1.02
13	14.40	14.25	14.70	14.90	5.15	7.76

Table 1. The three sets of initial mean $\boldsymbol{\mu}_k^0$ selected for Fig. 1.
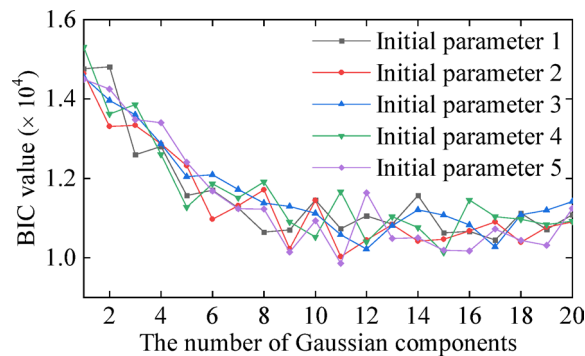
Fig. 2. BIC curves under different initial parameters. When the number of Gaussian components is 4, the initial weight of the k-th Gaussian component is set as $\alpha_k^0 = 1/4$ in each set of initial parameters, the initial covariance matrix of the k-th Gaussian component is $\boldsymbol{C}_k^0 = \begin{bmatrix} 86.57 & 1.89 \\ 1.89 & 97.09 \end{bmatrix}$, the initial mean of the k-th Gaussian component is listed in Table 2.

k-th Gaussian component	The 1st set		The 2nd set		The 3rd set		The 4th set		The 5th set	
	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
1	14.73	-0.10	-7.62	-9.41	-3.48	3.65	14.93	16.89	5.93	4.99
2	9.00	10.03	14.88	13.85	-7.97	-9.60	4.34	5.23	15.88	1.07
3	4.39	5.09	13.45	13.92	18.44	-9.17	18.44	-9.17	18.46	-0.82
4	6.08	11.21	18.54	-9.06	-6.22	11.28	-8.10	-8.65	-0.35	-6.82

Table 2. The selected five sets of initial mean $\boldsymbol{\mu}_k^0$ when the number of Gaussian components is 4.

method, are then used as the initial parameters for the GMM. This enhancement improves the modeling accuracy of the conventional GMM.

The sequential iteration method can be summarized in the following detailed steps[67]:

Step 1: Set the Gaussian cluster index k to 1, initialize the covariance parameter σ^2 of the dataset $\boldsymbol{X}$ using Eq. (9)[63] and randomly choose a sample point $\widehat{\boldsymbol{\mu}}_k$ from the training samples as the initial point. Let $\widehat{\boldsymbol{C}}_k = \sigma^2 \cdot \boldsymbol{I}$, and $\boldsymbol{I}$ represents the d-dimensional identity matrix.

$$\sigma^2 = \frac{1}{10d} trace\left(\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{x})(\boldsymbol{x}_i - \overline{x})^{\mathrm{T}}\right), \overline{x} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i \tag{9}$$

Step 2: Calculate the Mahalanobis distance between $\boldsymbol{x}_i\ (i = 1, \ldots, n)$ and the arbitrarily chosen point $\widehat{\boldsymbol{\mu}}_k \in \boldsymbol{X}$ using Eq. (10).

$$D\left(\boldsymbol{x}_i, \widehat{\boldsymbol{\mu}}_k\right) = \sqrt{\left(\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}}_k\right)^{\mathrm{T}} \widehat{\boldsymbol{C}}_k^{-1} \left(\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}}_k\right)} \tag{10}$$

If the Mahalanobis distance is below 3, assign the $\boldsymbol{x}_i$ into the sample set $\boldsymbol{X}_s^k$.

Step 3: Calculate the mean $\widetilde{\boldsymbol{\mu}}_k$ and variance $\widetilde{\boldsymbol{C}}_k$ of the sample set $\boldsymbol{X}_s^k$ using Eqs. (11)-(12),

$$\widetilde{\boldsymbol{\mu}}_k = \frac{1}{n_k} \sum\nolimits_{\boldsymbol{x}_i \in \boldsymbol{X}_s^k} \boldsymbol{x}_i \tag{11}$$

$$\widetilde{\boldsymbol{C}}_k = \frac{1}{n_k - 1} \sum\nolimits_{\boldsymbol{x}_i \in \boldsymbol{X}_s^k} \left(\boldsymbol{x}_i - \boldsymbol{\mu}_k\right) \left(\boldsymbol{x}_i - \boldsymbol{\mu}_k\right)^T \tag{12}$$

where n_k represents the number of samples in $\boldsymbol{X}_s^k$.

Step 4: Let $\widehat{\boldsymbol{\mu}}_k = \widetilde{\boldsymbol{\mu}}_k$, $\widehat{\boldsymbol{C}}_k = \widetilde{\boldsymbol{C}}_k$, and repeat Steps (2)-(3) until:

$$\| \widetilde{\boldsymbol{\mu}}_k - \widehat{\boldsymbol{\mu}}_k \| < \varepsilon \tag{13}$$

where ϵ is a small enough value, ϵ is set to 10^{-6} in this paper.

Step 5: Determine the sample set $\boldsymbol{X}_s^k$, compute the mean $\widehat{\boldsymbol{\mu}}_k$, covariance matrix $\widehat{\boldsymbol{C}}_k$, and weight coefficient $\widehat{\alpha}_k = \frac{n_k}{n}$ of $\boldsymbol{X}_s^k$. Then, from the dataset $\boldsymbol{X}$, remove the sample points $\boldsymbol{x}_i$ that have been allocated to $\boldsymbol{X}_s^k$;

Step 6: Increment k by 1 ($k = k + 1$).

Step 7: Repeat Steps (2)-(6) until all the data is assigned to the corresponding $\boldsymbol{X}_s^k$. Finally, the $\widehat{\boldsymbol{\mu}}_k$, $\widehat{\boldsymbol{C}}_k$, and $\widehat{\alpha}_k$ obtained from each sample set $\boldsymbol{X}_s^k$ are considered as the initial mean, covariance matrix, and weight in GMM.

To better visualize the stages of the proposed algorithm, the entire procedure for implementing the sequential iteration operation is illustrated in Fig. 3. The improved GMM eliminates the need for manual determination of the number of Gaussian components, as the sequential iteration process can automatically determine this value. Moreover, the model parameters obtained by the sequential iteration method is used as the initial model parameters of the EM algorithm, which can improve the computational efficiency and data fitting accuracy of the conventional GMM method.
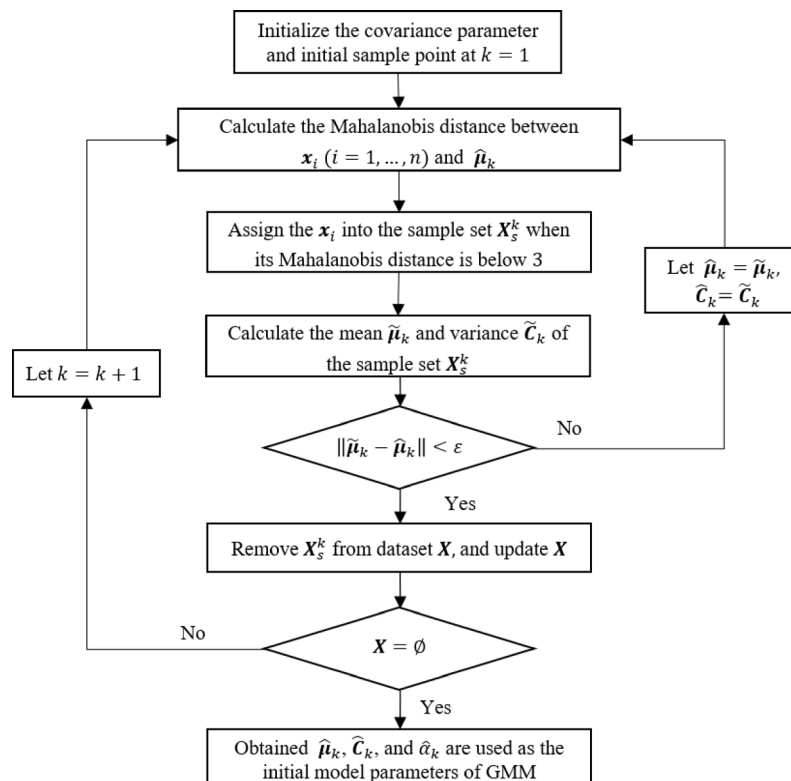
Fig. 3. Procedure for implementing the sequential iteration operation.

Implementation process of the proposed GMM damage detection algorithm

After enhancing the conventional GMM with a sequential iteration method, we utilize the improved GMM approach for detecting damage in changing environmental conditions. Initially, continuous monitoring data is collected in a healthy state under the changing environmental conditions. From these data, damage features are extracted, forming a baseline training sample set denoted as X. Next, a sequential iteration approach is employed to establish the initial model parameters of the GMM.

. The EM algorithm is then utilized to construct a baseline GMM based on the data in X. Furthermore, the MSD between the damage feature and each Gaussian component is calculated as follows:

$$MSD_k\left(x\right) = \left(x - \mu_k\right)^T C_k^{-1}\left(x - \mu_k\right) \tag{14}$$

The MSD has been shown to be an effective damage indicator that is insensitive to changing environments[37]. However, it is only capable of addressing linear EOV and remains sensitive to nonlinear EOV effects. To tackle this issue, Figueiredo et al. proposed the MMSD as a novelty index for damage detection under changing environments[38,39]. Similar studies utilizing the MMSD for this purpose can also be found in the literatures[54,68,69]. The minimum MSD is calculated as follows:

$$MMSD\left(x\right) = \min\left[MSD_1\left(x\right), \; \cdots, \; MSD_K\left(x\right)\right] \tag{15}$$

The insensitivity of the MMSD to nonlinear environmental effects stems from the division of nonlinearly related data into several categories using GMM. Compared to the entire dataset, each localized category exhibits a stronger linear correlation after clustering. Therefore, the MSD can effectively eliminate the linear environmental effects in these local datasets. In practice, the MMSD is calculated as the MSD of a sample point within its respective category. By separating the linear EOV effects within each category, the overall nonlinear environmental effects on the entire dataset are mitigated. This makes the MMSD insensitive to nonlinear EOV effects. In this paper, the MMSD is also used as the novelty index (NI) for damage detection under changing environments.

Furthermore, the MMSD values obtained during the training phase are used to compute a threshold for damage detection. When damage occurs in a structure, the feature data associated with the damaged state will deviate from the original distribution characteristics. Consequently, the MMSD will exceed the threshold during the testing phase, clearly indicating the presence of damage. To provide a comprehensive understanding of the algorithm's stages, the complete implementation procedure for the enhanced GMM method in damage detection is illustrated in Fig. 4.

To quantify the effectiveness of damage detection, this study employs two metrics: the false positive rate (FPR) and the false negative rate (FNR). The two metrics originate from the field of machine learning, where a confusion matrix table enables visualization of the performance of a machine learning algorithm based on some criteria, including true positive (TP), false positive (FP), false negative (FT), and true negative (TN). In the context of SHM, the terms "positive" and "negative" refer to the damaged and undamaged states, respectively[21]. The TN means that the structure is in undamaged condition and the method could correctly detect the undamaged state (i.e., the novelty indexes are below the threshold in the undamaged state); the FN means that the structure suffered from damage, but the method cannot correctly alarm the occurrence of damage (i.e., the novelty indexes are below the threshold in the damaged state); the FP means that the structure is in undamaged condition, but the method cannot correctly detect the undamaged state (i.e., the novelty indexes are over the threshold in undamaged state); the TP means that the structure suffered from damage and the method could correctly alarm
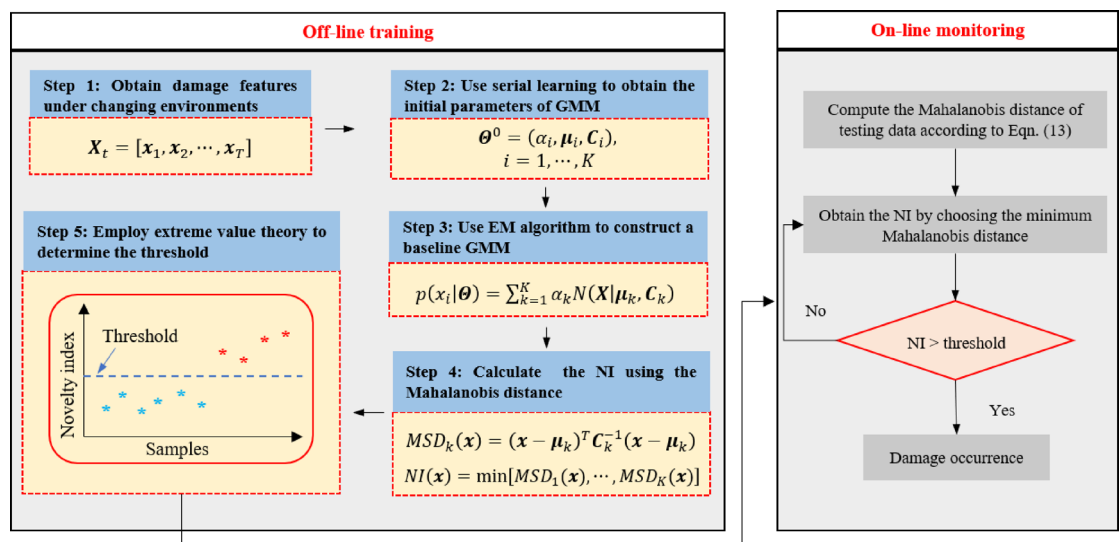
Fig. 4. The procedure of improving the GMM method for damage detection.

the occurrence of damage (i.e., the novelty indexes are over the threshold in damaged state). Table 3 depicts the confusion matrix, where two metrics, including FPR and FNR, are defined as follows.

$$FPR = \frac{N_{FP}}{N_{FP} + N_{TN}} \times 100\% \tag{16}$$

$$FNR = \frac{N_{FN}}{N_{FN} + N_{TP}} \times 100\% \tag{17}$$

where N_{TN}, N_{FN}, N_{FP} and N_{TP} represent the number of *TN*, *FN*, *FP* and *TP* sample points, respectively.

Threshold limit determination

The estimation of the threshold plays a crucial role in improved GMM for early damage detection. This estimation is typically derived from the probabilistic characteristics of MMSD during the training phase. One commonly used approach is to employ a standard confidence interval based on the central limit theory. For instance, a 95% confidence interval for MMSD is often utilized, assuming that MMSD during the training stage follows a normal distribution[70]. Nevertheless, relying solely on the standard confidence interval proves inadequate when MMSDs exhibit a non-normal or heavy-tailed distribution. To determine an appropriate threshold, the extreme value statistics is used in the improved GMM method. Extreme value statistics is a methodology employed to analyze exceptional events to identify extreme values within a given probability distribution. This approach finds extensive application in diverse fields such as structural engineering, finance, earth sciences, and geological engineering.

Extreme value analysis proves invaluable in establishing control limits that mark a failure point, beyond which a failure or end-of-life event is likely to occur. The generalized extreme value distribution (GEV) comprises three main forms: the Gumbel (Type I) distribution, the Fréchet (Type II) distribution, and the Weibull (Type III) distribution, these three types of extreme value distributions can effectively model extreme events across a diverse range of datasets[71].

Gumbel distribution:

$$H_1(x; \mu, \sigma) = e^{-e^{-\frac{x-\mu}{\sigma}}}, \quad -\infty < x < \infty \tag{18}$$

Fréchet distribution:

$$H_2(x; \mu, \sigma, \alpha) = \begin{cases} 0 & x \leq \mu, \\ e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}} & x > \mu, \end{cases} \quad \alpha > 0 \tag{19}$$

Weibull distribution:

$$H_3(x; \mu, \sigma, \alpha) = \begin{cases} e^{-\left(-\frac{x-\mu}{\sigma}\right)^{\alpha}} & x \leq \mu, \\ 1 & x > \mu, \end{cases} \quad \alpha > 0 \tag{20}$$

where μ is the location parameter, σ is the scale parameter, and α is the shape parameter. From a mathematical standpoint, the three types of extreme value distributions can be unified and expressed in a common form.

$$H(x; \mu, \sigma, \xi) = \exp\left\{-\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, \quad 1 + \xi \frac{x-\mu}{\sigma} > 0 \tag{21}$$

where ξ represents the shape parameter in the GEV distribution, where its value determines the type of extreme value distribution[71]. The threshold limit determination based on the GEV distribution model contains three main steps, which includes[72]: the selection of one of the extreme value distributions, the estimation of unknown parameters of selected extreme value distribution, and the determination of its extreme quantile. The extreme quantile of the distribution is estimated at a given significance level by inverting Eq. (21), which yields,

$$x_p = \begin{cases} \mu + \frac{\sigma}{\xi}\left[(-\log(p))^{-\xi} - 1\right] & \xi \neq 0 \\ \mu - \sigma \, log(-\log(p)) & \xi = 0 \end{cases} \tag{22}$$

In this paper, a confidence level of 95% is chosen, treating the tail probability of 5% in the cumulative probability function.

Detected results	Actual state	
	Undamaged	Damage
Undamaged	TN	FN
Damage	FP	TP

Table 3. The confusion matrix for evaluating the performance of the damage detection method.

Contributions

The main contributions and novelty of this research can be summarized as:

 (i) Development of an innovative damage detection method: This research introduces a novel method specifically designed for long-term SHM under varying environmental conditions. The key innovation lies in the integration of sequential iteration and GMM. The advantages of this method are: (1) It is capable of automatically selecting the number of components, and it does not require careful initialization; (2) Unlike the density peak clustering-based GMM[58] it does not require to empirically determine the model parameters; (3) Compared to the BIC-based method for determining the number of components, the proposed method is more stable and reliable. Moreover, this approach effectively addresses environmental variability, including nonlinear effects, which most traditional normalization techniques (like MSD, PCA, and factor analysis) can only handle in a linear context.
 (ii) Determination of threshold using the GEV distribution method. Traditional threshold determination method is to employ a standard confidence interval based on the central limit theory, which assumes that the damage index follows a normal distribution. However, relying solely on the standard confidence interval proves inadequate when MMSDs exhibit a non-normal or heavy-tailed distribution. Therefore, the proposed method uses the extreme value distribution to determine an appropriate threshold, which ensures more accurate threshold estimation, reducing the likelihood of false alarms and improving the reliability of damage detection.

## Real-world application 1: a wooden truss Bridge

In this section, the experimental data of a wooden truss bridge are used to verify the effectiveness of the proposed GMM method. The wooden bridge is a laboratory truss structure monitored by Prof. Kullaa for several days under a changing environmental condition[73]. The monitoring system for the structure is shown in Fig. 5, in which an electrodynamic shaker is applied on the wooden bridge to produce a random white noise excitation and fifteen accelerometers are deployed at different locations to measure the acceleration responses. Moreover, modal frequencies and mode shapes of the structure were identified from the vibration measurements based on the stochastic subspace method. Because only modal frequencies of the seven modes were made available to researchers, the seven frequencies are used as the main damage-sensitive features for damage detection.

To simulate various damage scenarios, different sizes of masses were attached to the wooden bridge, as detailed in Table 4. The total mass of the structure was 36 kg, with the heaviest mass added to the structure weighing 193.7 g. The monitoring process involved test measurements of both undamaged and damaged conditions of the bridge. The measurements 1-2000 were taken in undamaged state under varying environmental conditions, while the measurements 2001–2019, 2020–2042, 2043–2065, 2066–2091, and 2092–2114 were performed on the damaged state of the structure at different damage levels.

Figure 6 shows the changes of natural frequencies f_1, f_2, f_6 and f_7 over time, with the vertical dashed line indicating the moment of damage occurrence. It is apparent that environmental variations significantly impact the natural frequencies throughout the monitoring period, making it challenging to distinguish whether the frequency changes stem from structural damage or changing environments.

### Comparison of the proposed method with other methods

In the context of mitigating environmental influences on damage detection, cointegration[34] and PCA[28] are widely used techniques. This paper compares the proposed method with these two damage detection methods. Additionally, since the proposed method improves upon the GMM, we also compare the damage detection results with those of the conventional GMM to highlight the superiority of our approach. Using PCA, cointegration, and both conventional and proposed GMM methods, we implement damage detection for a Wooden Bridge during both training and test phases. During offline learning in the training period, 1800 observations under normal conditions are utilized as the training dataset. For the test phase, the remaining observations from normal conditions, alongside all observations from the damaged state, are utilized as the test dataset.

(a) (b)
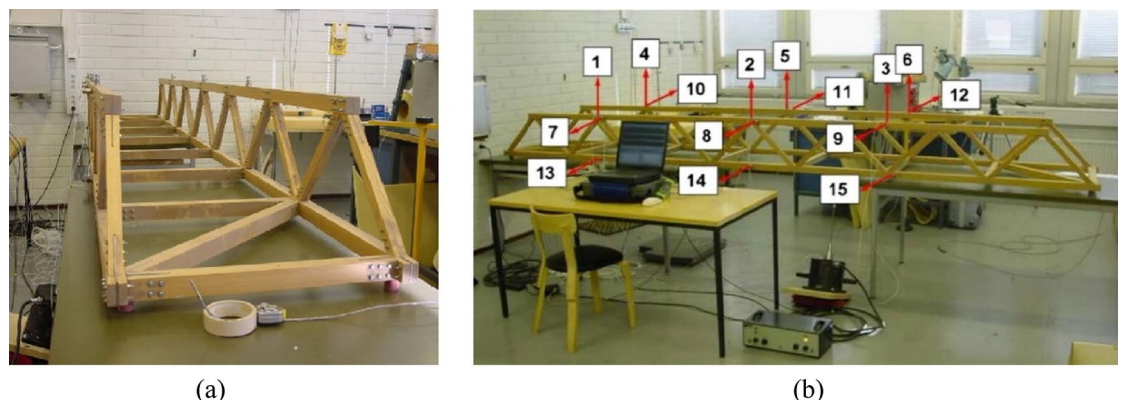
Fig. 5. (**a**) The Wooden Bridge, (**b**) the locations of the acceleration sensors and the electro-dynamic shaker[73].

| Damage case	Measurements	Added mass (g)
Case 1	[2001, 2019]	23.5
Case 2	[2020, 2042]	47
Case 3	[2043, 2065]	70.5
Case 4	[2066, 2091]	123.2
Case 5	[2092, 2114]	193.7

Table 4. Damage cases of the wooden truss bridge.
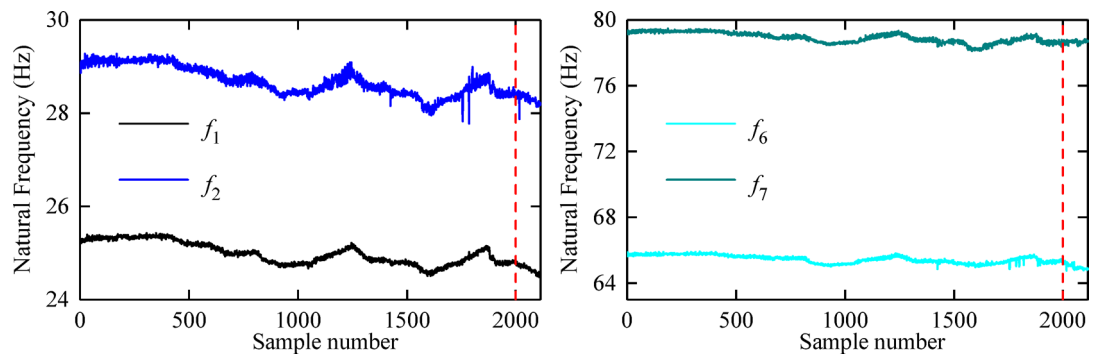
Fig. 6. Changes of the natural frequencies of the wooden bridge during the monitoring period (the occurrence of damage is denoted by the dashed vertical line).

Regarding the first comparison, it is necessary to calculate a stationary cointegration residual by linearly combining observed frequency data. The cointegration residual can be interpreted as a long-term stable equilibrium relationship between nonstationary frequencies. Once structural damage occurs, this equilibrium relationship will no longer be maintained, and the stationary residual will become nonstationary. Following the damage detection procedure of the cointegration method[34] the cointegration residual can be calculated by the Johansen test. Figure 7 displays the variation of cointegration residuals in the training and test phases, where the black scatter points represent the residuals in the training phase, the blue scatter points represent the undamaged residuals in the test phase, and the red points are the damaged residuals in the test phase.

Regarding the second comparison, the fundamental concept of PCA involves projecting the original data onto a vector space formed by the principal components (PCs) and then mapping it back to the original space while retaining a specific number of PCs. The error between the original and remapped data can be utilized to compute a novelty index for damage detection. Therefore, it is initially necessary to determine the number of PCs. This can be achieved by calculating the ratio of the sum of eigenvalues of consecutive principal components to the total sum of all eigenvalues and identifying the smallest integer with a ratio exceeding a predefined threshold. In this case, the threshold is set at 95%. Accordingly, the optimal number of principal components for the Wooden Bridge is determined to be 2, which accounted for 98.37% of the data variance. Based on the damage detection procedure of the PCA method[28] the novelty index can be obtained and shown in Fig. 8.

From Fig. 7, it can be seen that the cointegration residual is stationary in undamaged state, and almost all the residuals of the normal condition related to either the training or validation samples are within the control limit, indicating that environmental effects are effectively removed and no damage is detected in the structure in undamaged state. On the contrary, Fig. 8 shows that the evolution process of the PCA novelty indexes is nonstationary, and some of the undamaged novelty indexes are outside the control limit, indicating that some undamaged samples are misclassified as damaged samples. Furthermore, the poor performance of the cointegration and PCA methods emerges from the damaged state, where most of the residuals and novelty indexes are within the control limit, implying large false detection. Regardless of the control limit, it is observed that numerous cointegration residuals and novelty indexes of the damaged state are in the same scales as the undamaged state, which implies the low damage detectability of the cointegration and PCA methods.

In the third comparison, the conventional GMM method is employed for structural damage detection, which determines the number of Gaussian components using the BIC and estimates the model parameters using the EM algorithm. Figure 9 illustrates five BIC curves resulting from running the EM algorithm five times with different initial parameters. Similar to the numerical simulation results in Sect. 2, noticeable fluctuations are observed in the BIC curve. Furthermore, it is noted that when the number of Gaussian components is 9, the BIC value is relatively small. Opting for 9 Gaussian components, damage detection based on the conventional GMM method is conducted, as depicted in Fig. 10. In this analysis, the EM algorithm is run three times under varying initial parameters. The results indicate that due to the randomly selected initial parameters, the damage detection outcomes vary. This suggests that damage detection based on the conventional GMM method is prone to instability.

(a)　　　　　　　　　　　　(b)
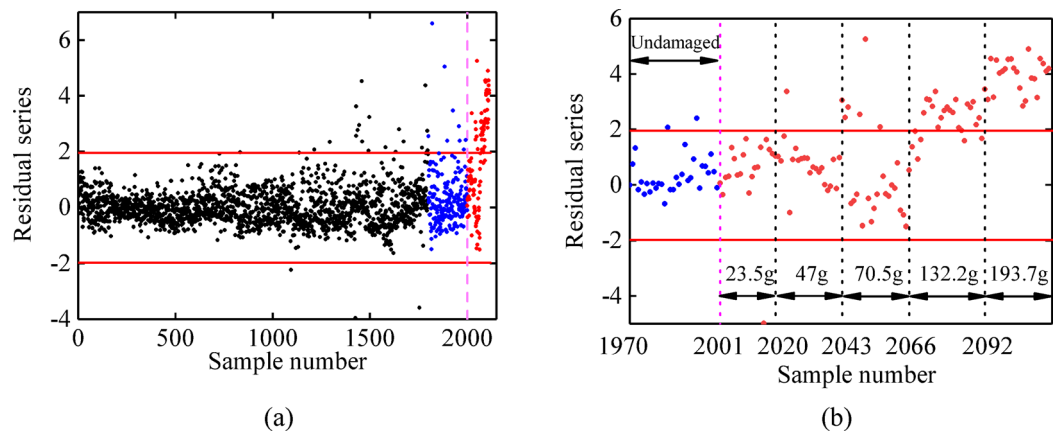
Fig. 7. Damage detection in the wooden bridge based on the cointegration method: (**a**) entire residuals and (**b**) partial residuals. (The black scatter points represent the residuals in the training phase, the blue scatter points represent the undamaged residuals in the test phase, and the red points are the damaged residuals in the test phase).
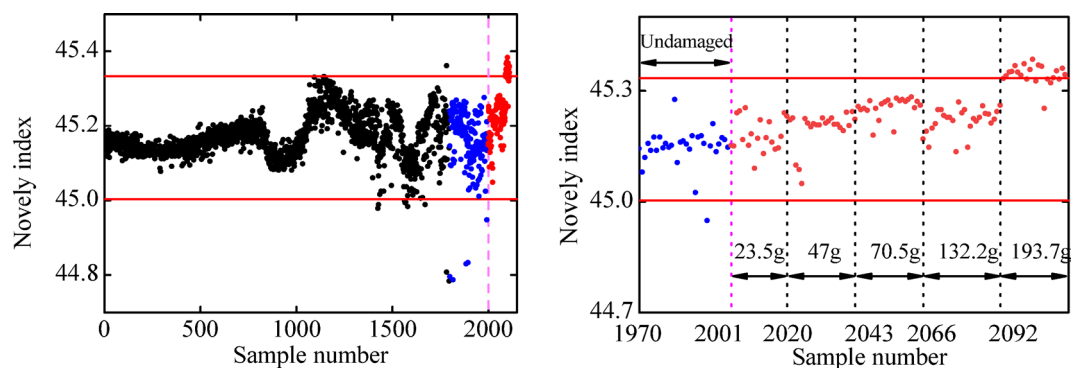
Fig. 8. Damage detection in the wooden bridge based on the PCA method: (**a**) entire novelty index and (**b**) partial novelty index. (The black points represent the novelty indexes in the training phase, the blue scatter points represent the undamaged novelty indexes in the test phase, and the red points are the damaged novelty indexes in the test phase).
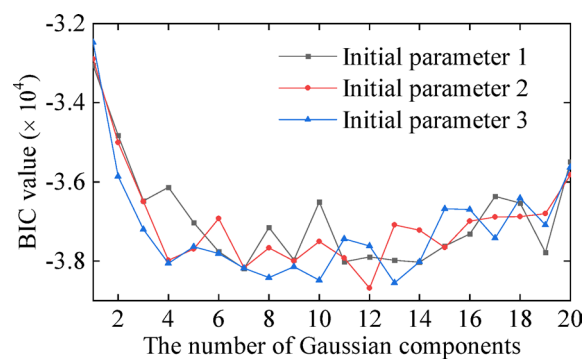
Fig. 9. BIC curves under three different initial states.
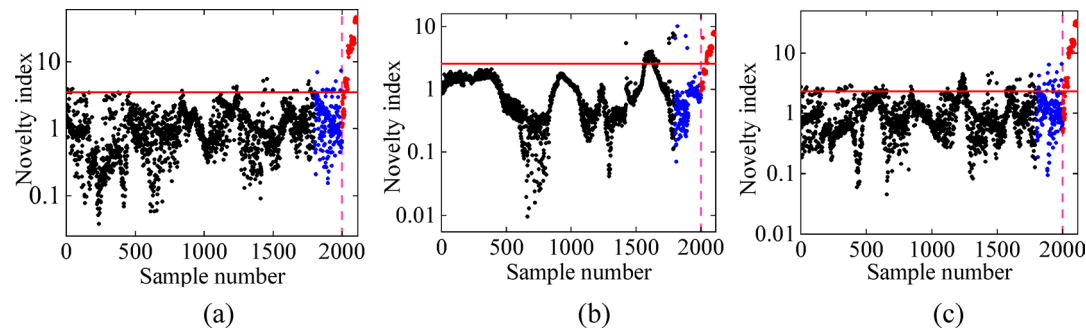
Fig. 10. Damage detection in the wooden bridge based on the conventional GMM method using 3 different sets of initial parameters. For each set of initial parameters, the initial mean is listed in Tables 5, 6 and 7, the initial weight of each Gaussian component is set as $\alpha_k^0 = 1/9$, the initial covariance matrix of each Gaussian component is set as

k-th Gaussian component	The 1st set of initial mean						
	f_1	f_2	f_3	f_4	f_5	f_6	f_7
1	25.16	28.92	39.28	52.93	63.89	65.57	79.13
2	25.27	29.00	39.37	53.04	64.00	65.70	79.22
3	24.62	28.11	38.89	52.37	63.77	65.14	78.13
4	24.79	28.39	38.98	52.47	63.64	65.19	78.71
5	24.79	28.41	38.99	52.46	63.66	65.19	78.71
6	24.79	28.44	39.04	52.57	63.81	65.36	78.61
7	25.31	29.08	39.45	53.12	64.04	65.81	79.26
8	25.28	29.01	39.38	53.03	63.98	65.70	79.23
9	24.76	28.38	39.01	52.45	63.65	65.24	78.54

Table 5. The 1st set of initial mean μ_k^0 when the number of Gaussian components is 9.

| *k*-th Gaussian component	The 2nd set of initial mean						
	f_1	f_2	f_3	f_4	f_5	f_6	f_7
1	25.34	29.08	39.47	53.10	64.02	65.79	79.29
2	25.19	28.99	39.28	52.94	63.81	65.57	79.23
3	24.82	28.49	39.02	52.59	63.83	65.36	78.73
4	25.00	28.72	39.16	52.81	63.94	65.55	78.93
5	25.28	29.02	39.43	53.07	64.00	65.78	79.23
6	24.90	28.55	39.12	52.66	63.83	65.44	78.81
7	24.86	28.58	39.05	52.60	63.69	65.28	78.87
8	24.85	28.45	39.11	52.69	63.87	65.43	78.75
9	24.99	28.77	39.16	52.76	63.73	65.43	79.06

Table 6. The 2nd set of initial mean μ_k^0 when the number of Gaussian components is 9.

$$C_k^0 = \begin{bmatrix} 0.06 & 0.07 & 0.05 & 0.06 & 0.03 & 0.05 & 0.07 \\ 0.07 & 0.10 & 0.06 & 0.08 & 0.04 & 0.07 & 0.09 \\ 0.05 & 0.06 & 0.04 & 0.05 & 0.03 & 0.04 & 0.06 \\ 0.06 & 0.08 & 0.05 & 0.07 & 0.04 & 0.06 & 0.08 \\ 0.03 & 0.04 & 0.03 & 0.04 & 0.02 & 0.04 & 0.04 \\ 0.05 & 0.07 & 0.04 & 0.06 & 0.04 & 0.06 & 0.07 \\ 0.07 & 0.09 & 0.06 & 0.08 & 0.04 & 0.07 & 0.10 \end{bmatrix}$$

The primary step in evaluating the performance of the proposed GMM-based method involves determining the initial GMM parameters through sequential iteration. For the process of the sequential iteration method, it begins with random selection of initial points $\widehat{\mu}_0$, which is here chosen as $\widehat{\mu}_0 = [24.80, 28.46, 39.01, 52.57, 63.79, 65.33, 78.64]^{\mathrm{T}}$. Subsequently, the mean, covariance matrix, and weight of each data set can be obtained by executing Steps 1–3 of the sequential iteration in Sect. 3.1. These

k-th Gaussian component	The 3rd set of initial mean						
	f_1	f_2	f_3	f_4	f_5	f_6	f_7
1	25.12	28.83	39.25	52.91	63.94	65.61	79.05
2	24.85	28.63	39.03	52.62	63.69	65.30	78.86
3	25.13	28.90	39.30	52.90	63.84	65.57	79.24
4	24.71	28.34	38.90	52.41	63.60	65.17	78.63
5	25.36	29.12	39.44	53.09	64.00	65.77	79.32
6	24.98	28.65	39.14	52.83	63.92	65.59	78.92
7	25.00	28.72	39.18	52.81	63.86	65.52	78.95
8	25.00	28.69	39.18	52.82	63.89	65.55	78.95
9	24.82	28.56	38.99	52.58	63.68	65.28	78.84

Table 7. The 3rd set of initial mean $\boldsymbol{\mu}_k^0$ when the number of Gaussian components is 9.
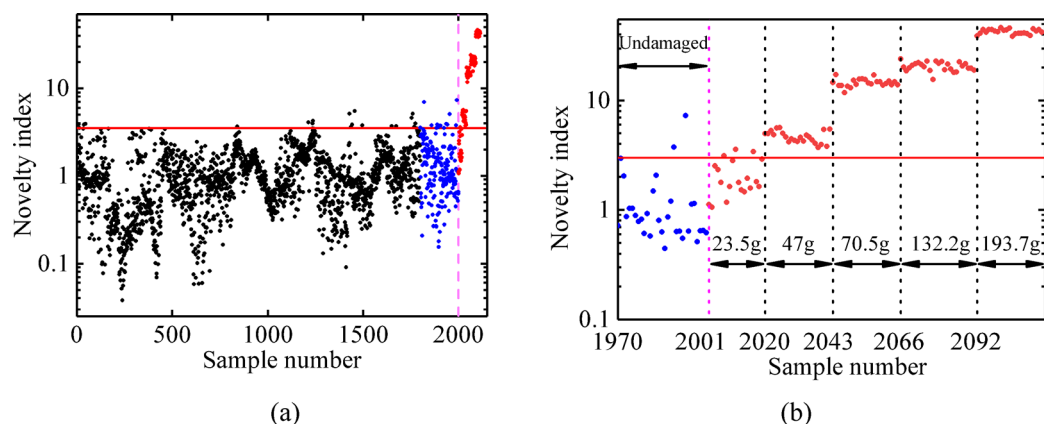
Fig. 11. Damage detection in the wooden bridge based on the proposed GMM method: (a) entire novelty index and (b) partial novelty index.

outputs serve as the initial parameters for establishing the GMM model. The next step involves calculating the MSD between the sampling points and each Gaussian component. The minimum distance is then chosen as the novelty index for damage detection. Figure 11 illustrates the results of early damage detection in the wooden truss bridge using the proposed GMM method. It is evident that the majority of novelty indexes related to the training samples fall below the threshold line, indicating a correct detection of the undamaged state with minimal false alarms. Moreover, the vast majority of novelty indexes for the validation data within samples 1800–2000 are also below the threshold. These observations confirm the efficacy of the proposed method in accurately detecting the undamaged state of the wooden bridge. In the case of novelty indexes related to the damaged state in samples 2001–2114, almost all of them exceed the threshold limit, demonstrating the high damage detectability of the proposed method. Nonetheless, it's worth noting that some observations in damage case 1 (23.5 g) fall below the threshold limit, which aligns with findings in references[21,51]. This behavior can be attributed to the relatively small mass, accounting for only 6.53×10^{-7} of the total structure mass, which causes a negligible deviation in natural frequencies from the undamaged cases. Regardless of the threshold line, it is observed that there is a clear discrepancy between the novelty indexes of the undamaged and damaged states. This outcome reinforces the robustness of the proposed method in providing discerning novelty indexes and achieving high damage detectability.

Table 8 presents the FPR and FNR calculated using cointegration, PCA, the conventional GMM, and the proposed GMM method. The FPR and FNR values for the conventional GMM are averaged from three results shown in Fig. 10. Regarding the FPR metric, the values for all three methods are relatively small, indicating minimal misclassifications under normal conditions. Concerning the FNR metric, the proposed GMM method significantly outperforms the other methods, suggesting a higher capability for accurate damage detection. The limitations of PCA and cointegration methods in accurate damage detection can be attributed to their linear nature. These methods attempt to derive a damage index by linearly combining different frequencies, assuming a strong linear relationship between natural frequencies. However, the natural frequencies of the wooden bridge exhibit inadequate linear correlation, leading to reduced accuracy in damage detection for PCA and cointegration methods. Additionally, the damage detection results obtained by the conventional GMM are not reliable due to its sensitivity to initial parameters. In contrast, the proposed GMM method divides the data into multiple datasets, treating each Gaussian component of the mixture as a cluster. This approach effectively decomposes nonlinearly related data into multiple locally linearly related datasets, transforming nonlinear problems into

	Cointegration	PCA	GMM	Proposed GMM method
FPR	1%	2%	6.86%	1.65%
FNR	61.40%	81.60%	27.78%	15.79%

Table 8. FPR and FNR obtained based on different damage detection methods.
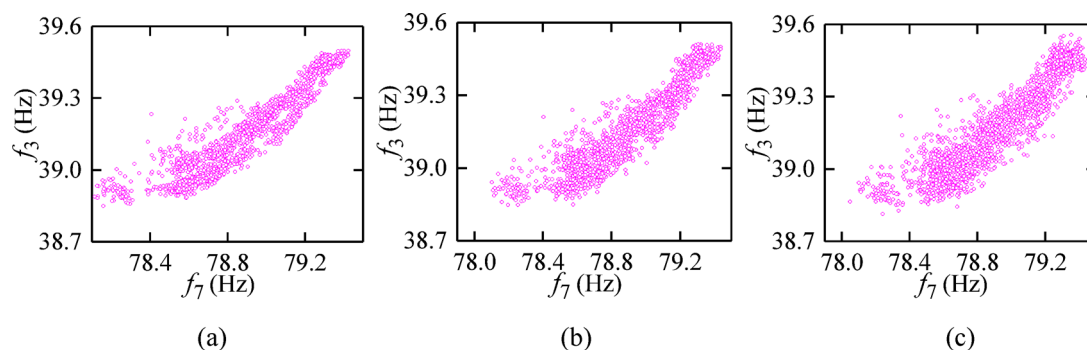
Fig. 12. The relation between f_3 and f_7 under (**a**) 0%, (**b**) 5% and (**c**) 10% noise levels.

linear ones. Consequently, the proposed method achieves more reliable damage detection, particularly when dealing with data exhibiting poor linear correlation, surpassing the capabilities of the cointegration and PCA methods.

Effects of measurement noise and initial points

Previous evaluations were conducted under laboratory conditions with minimal measurement noise, field measurements often encounter higher levels of noise in the recorded signals. The incorporation of Gaussian white noise into the original frequency data allows for a comprehensive assessment of the method's ability to handle real-world conditions with varying noise levels. Gaussian white noise was introduced into the original frequency data to examine the damage detection performance of the proposed GMM method using noisy data.

$$\boldsymbol{f}\left(t\right) = \widetilde{\boldsymbol{f}}\left(t\right) + E_p \mathcal{N}_{\text{noise}} \sigma\left(\widetilde{\boldsymbol{f}}\right) \tag{23}$$

where $\boldsymbol{f}\left(t\right)$ is the measured frequency data with noise, E_p is the noise level, $\mathcal{N}_{\text{noise}}$ is a Gauss-distributed random vector with zero mean and variance 1, and $\sigma\left(\widetilde{\boldsymbol{f}}\right)$ is the standard deviation of $\boldsymbol{f}$. Two levels of noise, namely 5% and 10%, were deliberately introduced into the original frequency data. The resulting correlations scatter plot between f_3 and f_7 at different noise levels is depicted in Fig. 12. Notably, as the noise level increases, the correlation between f_3 and f_7 exhibits a clear decrease, and the frequency sample points become noticeably more dispersed. This observation underscores the significant influence of measurement noise on the correlation between frequencies.

Following the procedure of the proposed GMM method in Sect. 3, the novelty indexes of the wooden bridge at different noise levels were obtained. Figures 13 and 14 present the results of damage detection at noise levels of 5% and 10%, respectively. It is evident that most novelty indexes in the undamaged state fall below the threshold limit, while some exceed it, possibly due to high noise levels. As the noise level increases, the misclassification rate also rises. However, even at a 10% noise level, most novelty indexes in the damaged state surpass the threshold value, indicating successful damage detection. Without considering the threshold limit, there is a discernible difference between the novelty indexes of the damaged and undamaged states, validating the high damage detectability of the proposed GMM method.

To demonstrate the stability and reliability of the proposed method, the damage identification process was repeated 20 times at various noise levels, given the random nature of measurement noise. The FNR and FPR metrics were computed for each of these 20 damage detections. Figure 12 displays the values of FNR and FPR for the 20 damage detections under different noise levels. From Fig. 15, several key observations can be made: (1) Similar to Fig. 14, as the noise level increases from 5 to 10%, both the FNR and FPR values also increase, indicating that higher noise levels can impact the accuracy of damage detection. (2) Despite the influence of measurement noise on the damage detection results, the FNR and FPR values remain relatively low, suggesting that the proposed GMM method exhibits excellent damage detectability even in the presence of high noise. (3) The amplitudes of FNR and FPR obtained from the 20 damage detections are remarkably stable, showing minimal fluctuations across different instances of damage detection. This stability underscores the high confidence in utilizing the proposed GMM method for damage detection. In summary, the repeated experiments at various noise levels illustrate the consistent and dependable performance of the proposed GMM
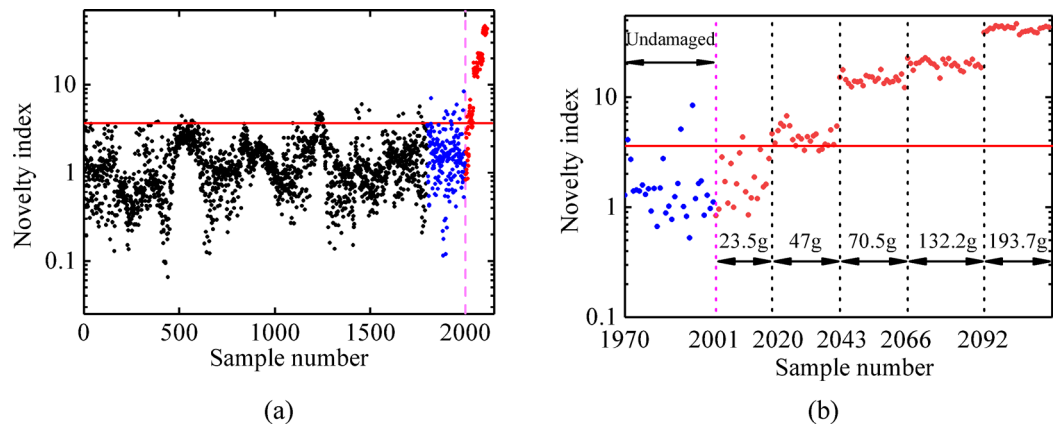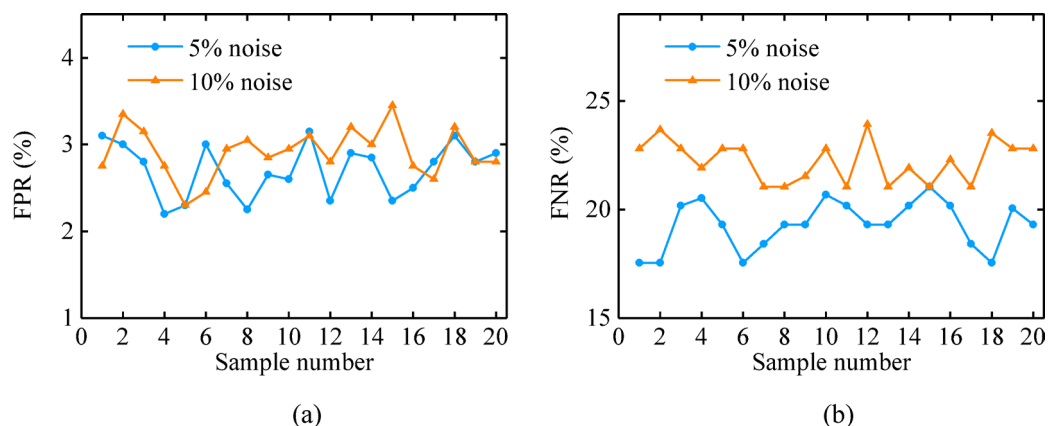
Fig. 13. Damage detection based on the proposed method under 5% noise level: (**a**) entire novelty index and (**b**) partial novelty index.
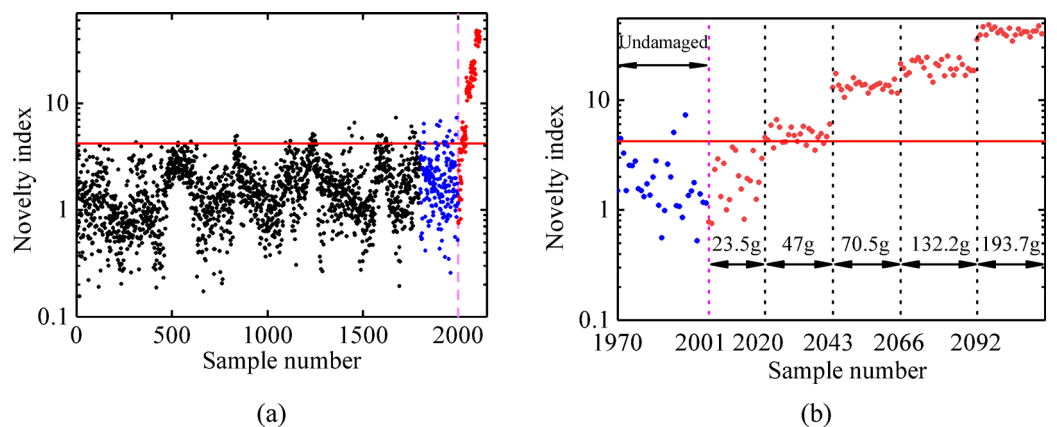
Fig. 14. Damage detection based on the proposed method under 10% noise level: (**a**) entire novelty index and (**b**) partial novelty index.
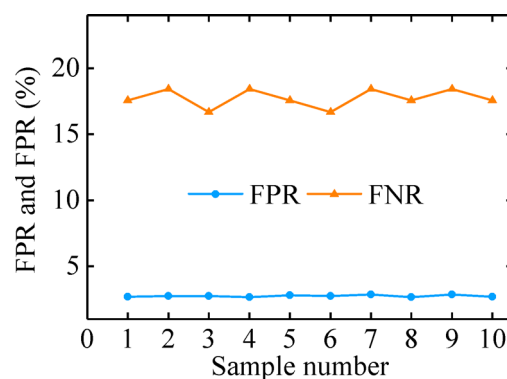

Fig. 15. FPR and FNR obtained in 20 repeated damage detection processes under different noise levels.

method in detecting damage, as evidenced by the relatively low FNR and FPR values and their stable nature throughout the experiments.

Regarding the proposed GMM method, it is initially necessary to specify an initial point needed for the sequential iteration process. As explained in Sect. 3.1, this initial point is randomly selected from the original dataset. To investigate the impact of randomly selected initial points on the method's performance, the damage

Initial point	The i-th selection of the initial point									
	1	2	3	4	5	6	7	8	9	10
f_1	24.74	25.15	24.78	25.14	24.87	24.89	24.89	25.05	24.60	25.01
f_2	28.35	29.12	28.13	28.76	28.47	28.75	27.77	28.81	28.07	28.74
f_3	39.04	39.35	39.23	39.32	39.14	39.11	39.26	39.22	38.87	39.19
f_4	52.56	52.93	52.58	53.03	52.66	52.68	52.72	52.87	52.39	52.80
f_5	63.67	63.86	63.33	64.12	63.99	63.70	63.59	63.99	63.74	63.96
f_6	65.29	65.64	64.84	65.75	65.45	65.40	65.05	65.58	65.24	65.54
f_7	78.66	79.29	78.42	79.12	78.55	79.07	78.75	79.03	78.21	78.86

Table 9. Ten randomly chosen initial points in 10 repeated damage detection processes.

Fig. 16. FPR and FNR obtained in 10 repeated damage detection processes using 10 different initial points.

detection process was repeated 10 times with 10 different random initial points. The values of these initial points are listed in Table 9, and Fig. 16 illustrates the corresponding FNR and FPR calculated for each damage detection. It is evident from the results that the FNR and FPR values are consistently low across all damage detections, with minimal fluctuations in their amplitudes. This finding indicates that the selection of different initial points has negligible effects on the proposed method. The proposed GMM method demonstrates reliability and robustness for SHM even under different initial points.

Real-world application 2: Z24 Bridge

To further validate the environmental robustness of the proposed GMM method for damage detection, a widely used benchmark structure, namely the Z24 bridge, is used in this section. Located in the province of Bern, Switzerland, the Z24 bridge was a three-span post-tensioned concrete bridge (Fig. 17) that was eventually demolished in 1998 to make way for a larger span railway bridge. Before its demolition, the bridge was equipped with a SHM system, which measured the vibration responses as well as environmental variables, such as acceleration, temperature, humidity, and wind characteristics. During the undamaged period of Z24 bridge, the SHM system continuously collected data on the bridge's behavior under normal operating conditions. Towards the end of the health monitoring period, a series of progressive damage scenarios were artificially applied to simulate different damage states. For more detailed information regarding the configuration of the SHM system and implemented vibration tests, interested readers can find comprehensive details in References[12].

An automatic modal analysis based on frequency domain decomposition was developed to extract the natural frequencies of the first four modes under varying environmental conditions. The dataset comprises 5652 observations, with the first 4848 observations representing the healthy state and the subsequent observations from 4849 to 5652 corresponding to the damaged state. In Fig. 18, the evolution of the four natural frequencies over time is presented, with the dashed vertical line indicating the moment of damage. The results reveal that changing environmental conditions have a significant impact on the natural frequencies of the Z24 bridge, and the frequency variations caused by environmental influences outweigh those caused by the damage. Particularly, an evident peak is observed around sample 2000, coinciding with a period of very low temperatures. Moreover, a bilinear relationship between frequency and temperature is observed, which further leads to a nonlinear relationship between the frequencies[44].

Damage detection

A comparative analysis was conducted to showcase the superiority of the proposed method over the PCA and cointegration methods. In the comparison, damage detection via different methods is carried out in the training and test phases. In the training phase, 80% of the undamaged observations of the modal frequencies were utilized

(a)　　　　　　　　　　　　　(b)
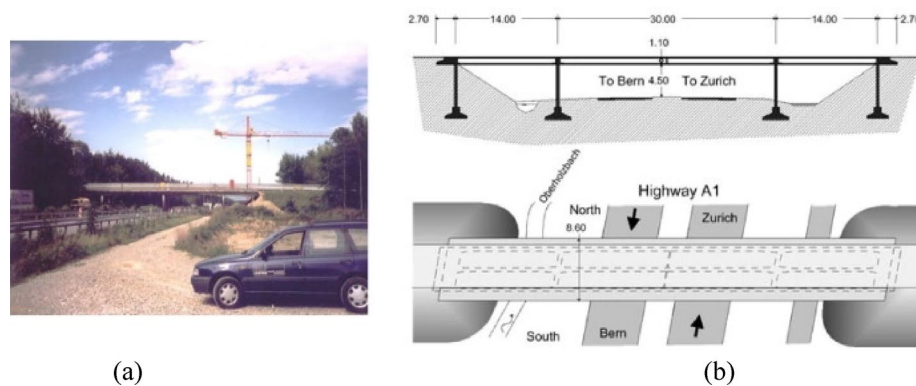
Fig. 17. (**a**) The Z24 Bridge, (**b**) the longitudinal section and top view of the bridge.
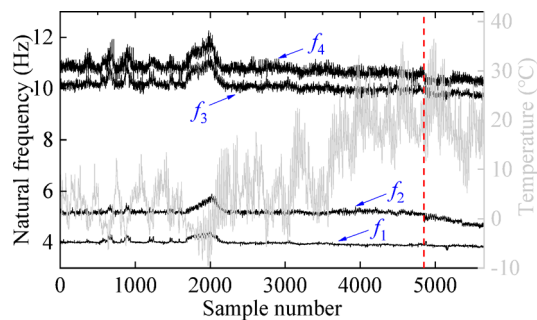

Fig. 18. Changes of the natural frequencies of Z24 Bridge during the monitoring period (the occurrence of damage is denoted by the dashed vertical line).

as the training dataset. Subsequently, the remaining 20% of the undamaged observations, along with all damaged observations, were considered as the test dataset.

In the first comparative study, the popular machine learning methods, Support Vector Regression (SVR) and Gaussian Process Regression (GPR), were employed to detect structural damage under varying EOCs. Initially, these methods involve establishing a regression relationship between temperature and modal frequency during the training phase. The trained model is then used to predict frequency based on the testing temperature data. Structural damage is detected by analyzing the prediction error. According to the above process, damage detection results can be obtained based on SVR and GPR methods, as shown in Figs. 19 and 20. It can be seen that when using the frequency f_1 as the output, both SVR and GPR methods produced a significant number of false alarms in damaged stage, indicating that damage was not effectively detected. Although the damage identification results are improved when using f_2 as the output, there are still a large number of undamaged samples that are mistakenly detected as damaged samples, indicating that there are a large number of false positive detection. This result shows that these two methods are inaccurate for damage detection of Z24 bridge.

In the second comparative study, by performing the damage detection procedures of the cointegration and PCA methods, the cointegration residuals and PCA novelty index were obtained (see Fig. 21). Observing Fig. 21, it is evident that almost all the residuals and novelty indexes in the undamaged state are below the threshold limit. This indicates that both methods yield only a few false alarms in the undamaged state. However, the cointegration residuals do not show a clear shift when the damage occurs (Fig. 21(a)), resulting in the cointegration method's failure to detect damage accurately and distinguish the damaged state from the undamaged state. Similarly, in the case of the PCA method (Fig. 21(b)), most of the novelty indexes in the damaged state do not exceed the threshold. This indicates a significant number of false negatives in detecting damage with the PCA method. Consequently, these observations highlight the poor damage detectability of both the cointegration and PCA methods under nonlinear environmental variability. Moreover, Fig. 22 illustrates the damage detection results of the conventional GMM using three sets of randomly selected initial parameters. It can be observed that due to its sensitivity to initial parameters, the detection results of the GMM vary each time, with only one result being reasonable. This highlights the instability of the conventional GMM method in detecting damage in the Z24 bridge.

$$C_k^0 = \begin{bmatrix} 0.01 & 0.01 & 0.02 & 0.02 \\ 0.01 & 0.02 & 0.02 & 0.02 \\ 0.02 & 0.02 & 0.07 & 0.06 \\ 0.02 & 0.02 & 0.06 & 0.08 \end{bmatrix}$$

(a)

(b)
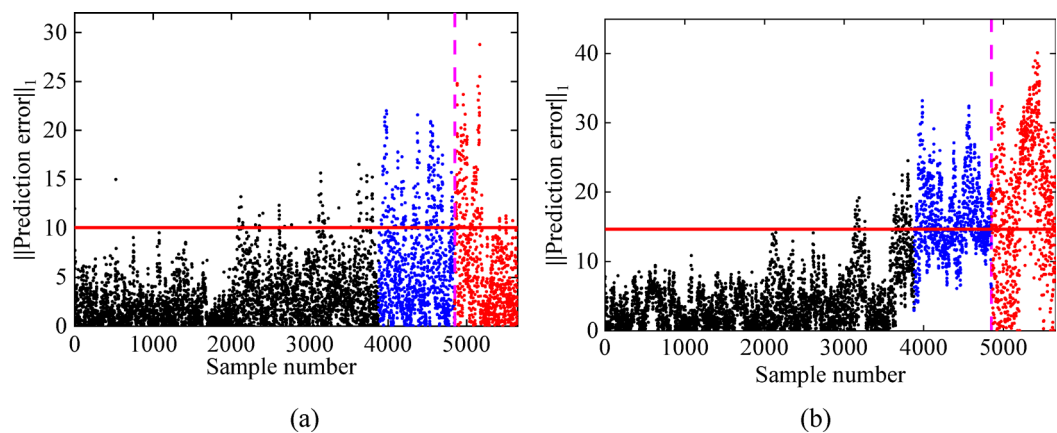
Fig. 19. Damage detection based on the SVR method using (**a**) temperature and f_1, (**b**) temperature and f_2 as the input and output, respectively. (The black scatter points represent the prediction errors in the training phase, the blue scatter points represent the undamaged prediction errors in the test phase, and the red points are the damaged prediction errors in the test phase).

(a)

(b)
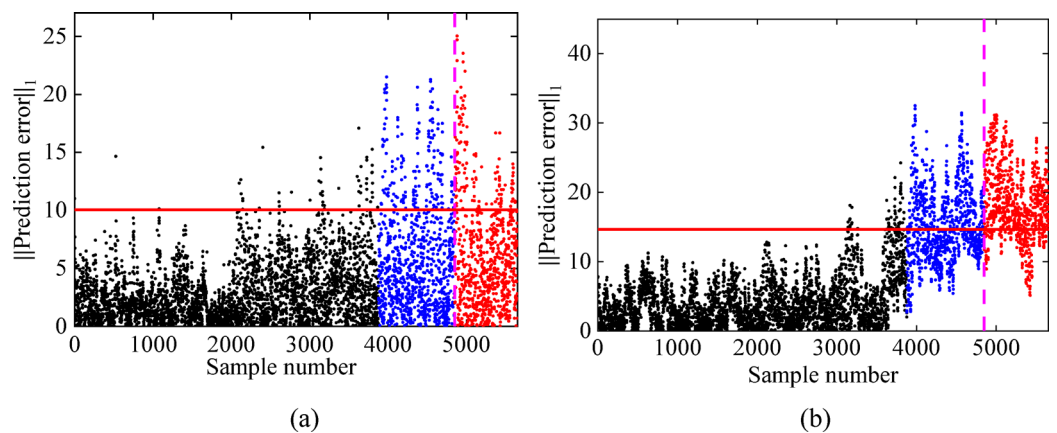
Fig. 20. Damage detection based on the GPR method using (**a**) temperature and f_1, (**b**) temperature and f_2 as the input and output, respectively. (The black scatter points represent the prediction errors in the training phase, the blue scatter points represent the undamaged prediction errors in the test phase, and the red points are the damaged prediction errors in the test phase).

(a)

(b)
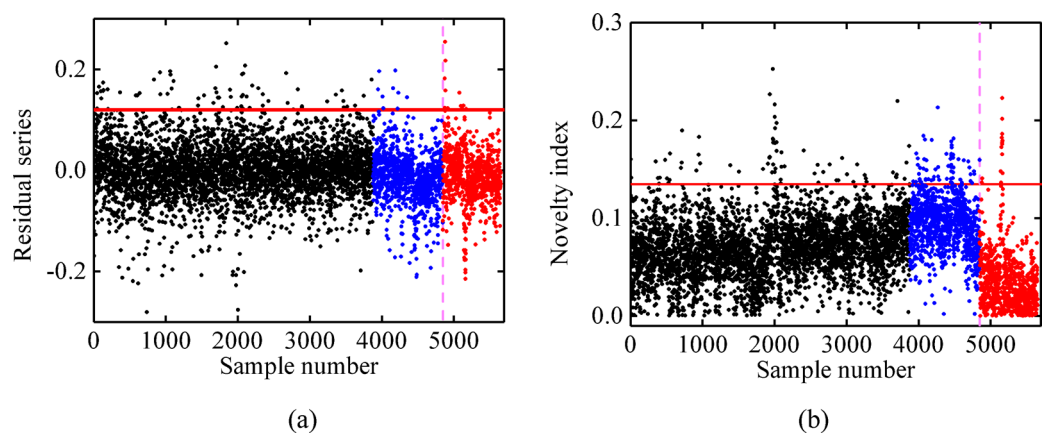
Fig. 21. Damage detection in the Z24 Bridge based on the conventional methods: (**a**) cointegration, (**b**) PCA. (The black scatter points represent the residuals in the training phase, the blue scatter points represent the undamaged residuals in the test phase, and the red points are the damaged residuals in the test phase).
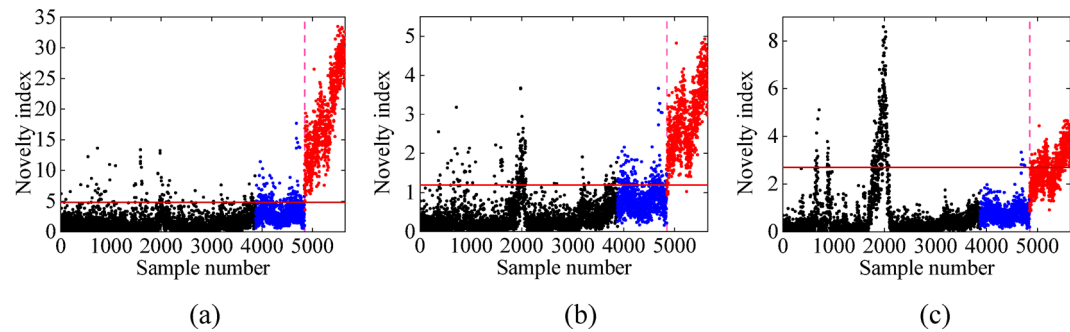
Fig. 22. Damage detection in the Z24 bridge based on the traditional GMM using 3 different sets of initial parameters. For each set of initial parameters, the three sets of initial mean are listed in Tables 10, 11 and 12, the initial weight of the k-th Gaussian component is set as $\alpha_k^0 = 1/2$, the initial covariance matrix of the k-th Gaussian component is set as,

k-th Gaussian component	The 1st set of initial mean			
	f_1	f_2	f_3	f_4
1	3.97	5.13	10.21	10.71
2	3.97	5.16	9.92	11.03

Table 10. The 1st set of initial mean μ_k^0 when the number of Gaussian components is 2.

k-th Gaussian component	The 2nd set of initial mean			
	f_1	f_2	f_3	f_4
1	4.03	5.16	10.12	10.91
2	4.02	5.20	10.12	10.79

Table 11. The 2nd set of initial mean μ_k^0 when the number of Gaussian components is 2.

k-th Gaussian component	The 3rd set of initial mean			
	f_1	f_2	f_3	f_4
1	3.96	5.10	10.16	10.68
2	3.93	5.15	10.10	10.66

Table 12. The 3rd set of initial mean μ_k^0 when the number of Gaussian components is 2.

Figure 23 displays the novelty index obtained using the proposed GMM method. From Fig. 23, it can be seen that in the undamaged state during the test phase, most of the novelty indexes fall below the threshold limit, suggesting that these observations correspond to the same structural condition as the observations in the training dataset (normal condition); a small number of undamaged observations are out of limit and this might be attributed to large measurement noise. In the damaged state, all the novelty indexes surpass the threshold limit, and the value of the novelty index increases with the severity of the damage. This indicates that the proposed method not only detects the occurrence of structural damage but also identifies the progression of damage in the structure accurately. The clear distinction between undamaged and damaged cases is evident regardless of the threshold value, confirming the high damage detectability of the proposed GMM method in accurately assessing the state of the Z24 bridge.

For further evaluation, Table 13 presents the FPR and FNR obtained using the cointegration, PCA, conventional and the proposed GMM methods. Although the cointegration and PCA methods demonstrate reasonable performances in terms of low FPR (minimizing the occurrence of false alarms), they suffer from a significant number of misclassifications with high FNR (failing to detect actual damage). This indicates that these two methods are not adequately suitable for effectively detecting structural damage in the Z24 bridge. On the other hand, the proposed GMM method exhibits the lowest rate of misclassification and triggers only a small number of false alarms. Hence, the comparison of the results of damage detection shown in Table 13 reveals that
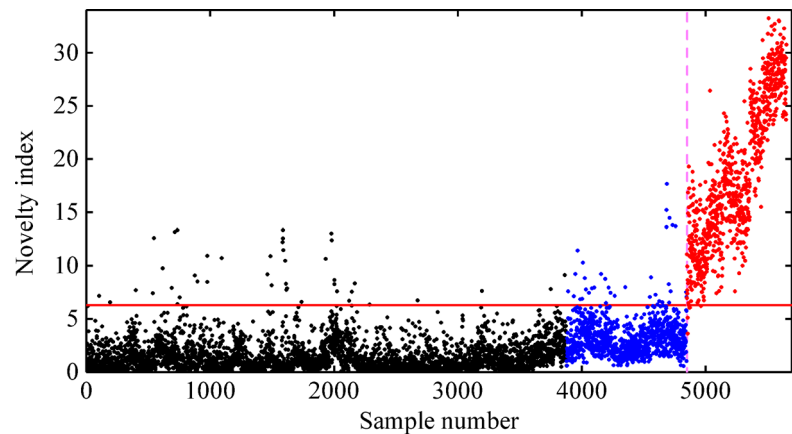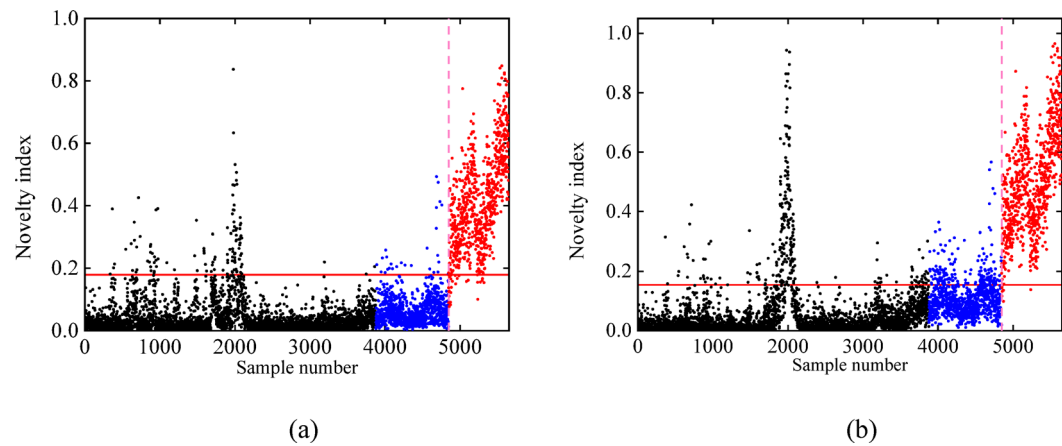
Fig. 23. Damage detection in the Z24 Bridge based on the proposed GMM method. (The black scatter points represent the residuals in the training phase, the blue scatter points represent the undamaged residuals in the test phase, and the red points are the damaged residuals in the test phase).

	Cointegration	PCA	GMM	Proposed GMM method
FPR	1.59%	2.74%	5.61%	1.63%
FNR	98.63%	98.26%	17.20%	0.75%

Table 13. FPR and FNR obtained based on different damage detection methods.

(a)

(b)

Fig. 24. Damage detection in the Z24 Bridge based on (**a**) the SI-K-means method and (**b**) the SI-FCM method.

the proposed GMM method outperforms the other methods in terms of having smaller errors and providing higher damage detectability.

In the following, the performance of the proposed GMM-based sequential iteration (SI) method is compared with clustering techniques based on the K-means and Fuzzy C-means (FCM) algorithms. For this comparison, the SI method is used to initialize the parameters of these algorithms. Damage detection results using the SI-K-means and SI-FCM techniques are presented in Fig. 24(a)–(b). It is evident that both SI-K-means and SI-FCM techniques produce a significant number of false alarms during the undamaged stage, resulting in ineffective damage detection. In contrast, the proposed method shows almost no false alarms, as observed in the Fig. 23. This result may be attributed to the fact that K-means and FCM assume clusters to be spherical and of equal size, which is not ideal for data with varying shapes or sizes. On the other hand, GMM can handle elliptical clusters and accommodates different covariance structures, allowing it to better model the Z24 frequency data, which is not spherically distributed. As a result, the GMM-based method outperforms K-means and FCM in terms of clustering accuracy, leading to more reliable damage detection outcomes.
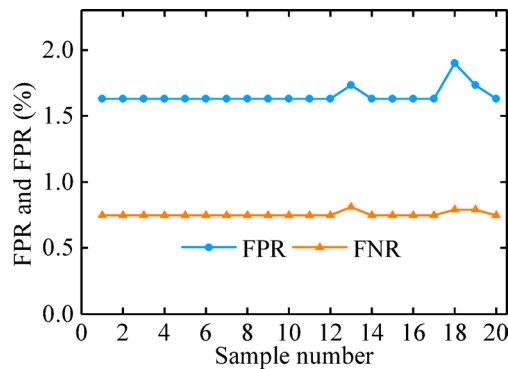
Fig. 25. FPR and FNR obtained in 20 repeated damage detection processes using 20 randomly chosen initial points.
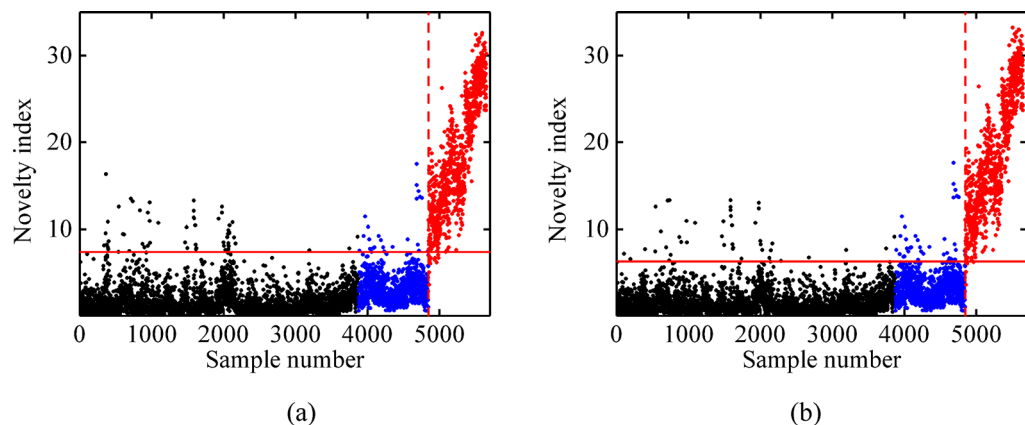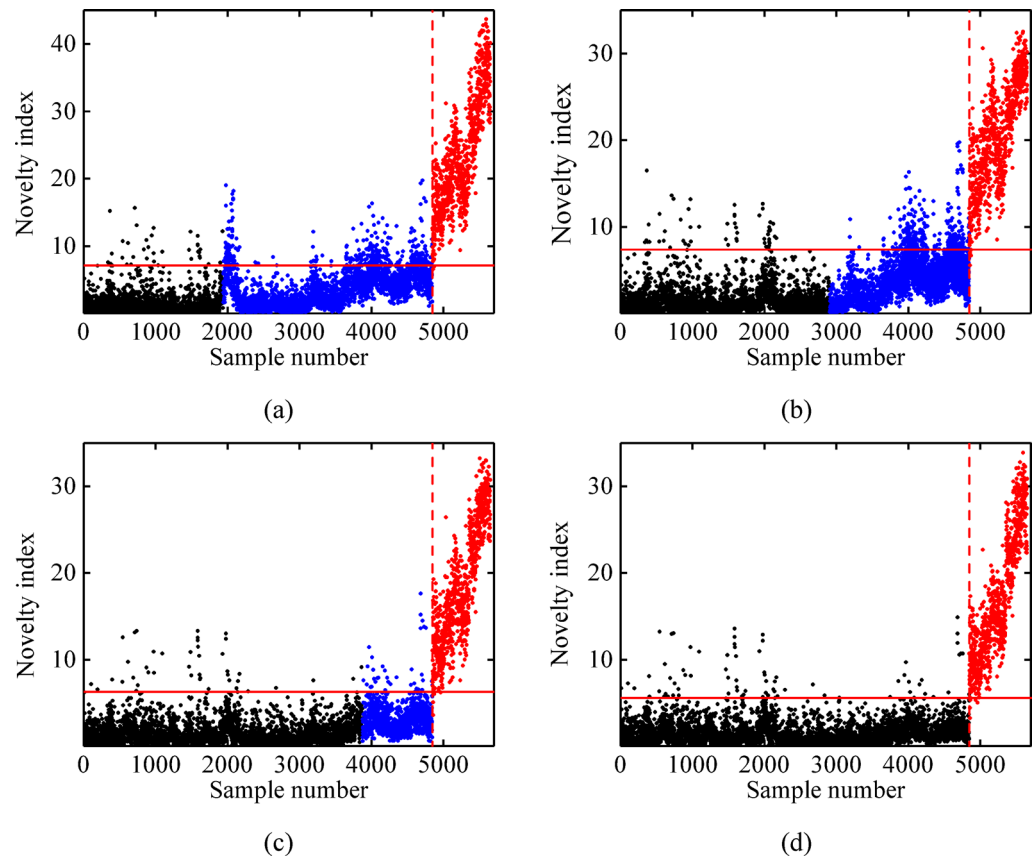

Fig. 26. Damage detection in the Z24 Bridge based on the proposed GMM method by using different initial points: (**a**) $\widehat{\boldsymbol{\mu}}_0 = [4.02, \ 5.04, \ 10.37, 11.38]^T$ and (**b**) $\widehat{\boldsymbol{\mu}}_0 = [3.91, \ 5.18, \ 9.73, 10.51]^T$.

Effects of different initial points and training data

To assess the impact of initial points on the damage detection of the proposed method for the Z24 bridge, we conducted 20 repeated damage detection processes using 20 randomly chosen initial points. Figure 25 presents the calculated FNR and FPR for each damage detection, while Fig. 26 illustrates the evolution of novelty index for two damage detections over time. It can be seen that the FPR and FNR obtained in each damage detection is very stable with less volatility, and almost no false alarms are observed in both undamaged and damage phases. The results clearly demonstrate the remarkable stability of the proposed GMM method for damage detection, affirming its insensitivity to the initial point selection. Furthermore, when progressive damage scenarios are introduced to the bridge, there is a notable increase in the novelty index values. This observation indicates the effective detection capability of the proposed method in identifying introduced damages, with the novelty index proving to be highly sensitive to structural damage while remaining unaffected by changing environmental conditions.

In all the previous analyses, the process of damage detection involved utilizing 80% of the observations of natural frequencies associated with the normal condition as the training data. In order to furtherly investigate the impact of the training dataset size on the performance of the proposed GMM damage detection method, new training datasets were created using 40%, 60%, and 100% of observations of the undamaged natural frequencies. These training sets comprised 1939, 2909, and 4848 learning samples, respectively, covering different wide ranges of environmental conditions under the normal condition. Following the procedure shown in Fig. 4 for the proposed GMM damage detection method, the novelty indexes are computed in training and test stages. Figure 27 shows the results of early damage detection in the Z24 Bridge via the proposed GMM method using different numbers of learning samples. Furthermore, Table 14 lists the FPR and FNR obtained under different training samples.

From Fig. 27; Table 14, it is evident that a significant number of misclassifications occur in the undamaged state when using 40% and 60% of the natural frequencies as training data. Particularly, a noticeable peak is observed around sample 2000 when utilizing 40% learning samples. Moreover, as the size of the learning samples increases, the misclassification rate decreases. Accurate and reliable damage detection can be achieved when using 80% and 100% of the observations from the normal condition. The observed phenomenon can be explained using Figs. 28 and 29, which depict the contour plots and probability distribution plots of f_1 based

Fig. 27. Damage detection in the Z24 Bridge by using the proposed GMM method under different learning samples: (**a**) 40%, (**b**) 60%, (**c**) 80% and (**d**) 100%.
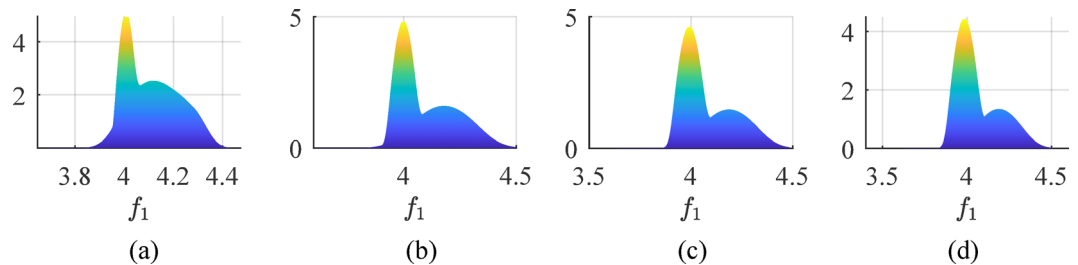
Learning samples	FPR	FNR
40%	7.45%	0.25%
60%	3.51%	0.50%
80%	1.63%	0.75%
100%	1.51%	0.87%

**Table 14**. FPR and FNR obtained via the proposed GMM method under different learning samples.
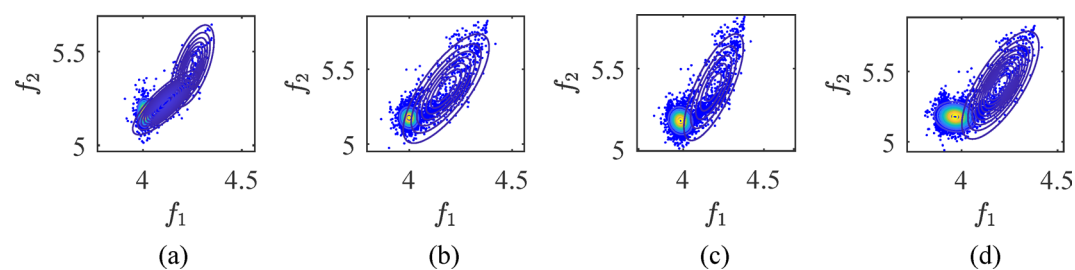
Fig. 28. Contour plot for f_1 and f_2 obtained by the proposed GMM method under different learning samples: (**a**) 40%, (**b**) 60%, (**c**) 80% and (**d**) 100%.

on the proposed method using different numbers of frequency data, denoted as f_1 and f_2. Notably, when only 40% of the undamaged data is used as the training data, the number of Gaussian components and distribution diagram notably differ from those obtained with other training data percentages. This discrepancy arises because 40% of training samples fail to encompass frequency data under low-temperature conditions, resulting

Fig. 29. Probability density of f_1 obtained by the proposed GMM method under different learning samples: (**a**) 40%, (**b**) 60%, (**c**) 80% and (**d**) 100%.
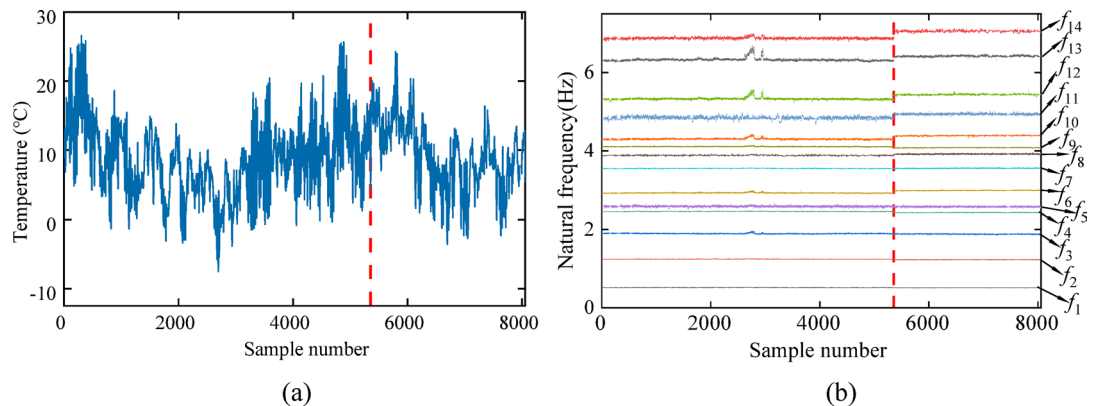

Fig. 30. Changes of (**a**) the surface temperature below the bridge deck and (**b**) the 14 natural frequency series over time. The period during which the retrofit is performed is indicated with a gray shading[74].

in an inaccurate probability distribution. Consequently, errors occur in the calculation of the MMSD, leading to inaccurate damage detection. This underscores the critical importance of employing a sufficient number of learning samples that cover a wide range of environmental variations to ensure the accuracy and reliability of the damage detection process.

Real-world application 3: KW51 Bridge

To further assess the effectiveness of the proposed GMM method on real structures, this section examines its application to the KW51 railway bridge. The KW51 bridge, a steel arch structure located on railway line L36N between Leuven and Brussels, Belgium, has been operational since 2003 [74]. In September 2018, a SHM system was installed to record various operational data, including acceleration, displacement, strain, temperature and relative humidity. From May 15 to September 27, 2019, the bridge underwent a retrofit to correct a construction error identified during inspections. This retrofit involved reinforcing the connections between the diagonals, arches, and bridge deck. Consequently, the collected data spans three distinct periods: pre-retrofit (7.5 months), during the retrofit (4.5 months), and post-retrofit (3.5 months). Figure 30 shows the variation of the model frequencies and temperature series before and after the retrofit. It can be seen that the fluctuations in natural frequencies are strongly affected by environmental conditions, especially the higher order modal frequencies. Around the 3000th sample, the stiffness of the asphalt layer on the bridge deck increased significantly due to frost caused by the lower temperature, resulting in a sharp increase in the natural frequency.

To evaluate the performance of various methods for damage identification, only the first 6 frequencies with low sensitivity to damage are selected for damage identification. Similar to the Z24 bridge analysis, 80% of the undamaged data was allocated for training, while the remaining 20% and all damaged data were used for testing. Figure 31 (a)-(d) presents the damage detection results for PCA, cointegration, conventional GMM, and the proposed GMM methods. A numerical comparison, summarized in Table 15, highlights the false positive rate (FPR) and false negative rate (FNR) for each approach. PCA and cointegration demonstrated the weakest performance, with high false negative rates, indicating their limited effectiveness in identifying damage in the KW51 bridge. The conventional GMM method showed improved results compared to PCA and cointegration; however, fluctuations in the novelty index during the undamaged phase suggest its limited ability to remove environmental influences. The proposed GMM method, by contrast, achieved the best performance, with the lowest error rates. During the undamaged phase, almost no novelty indices exceeded the control threshold, whereas all novelty indices during the damaged phase surpassed it. These findings confirm the proposed method's effectiveness in eliminating environmental effects and accurately detecting structural damage.
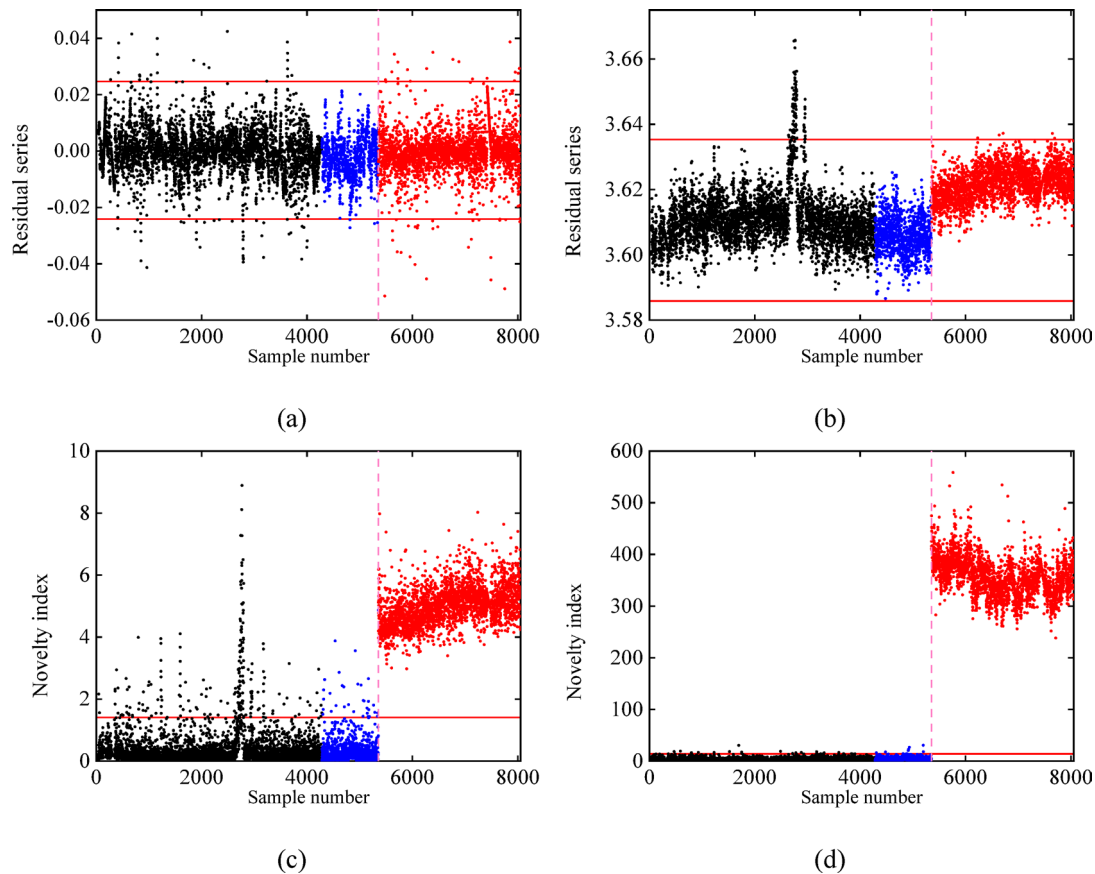
Fig. 31. Damage detection in the KW51 Bridge based on the conventional methods: (**a**) cointegration, (**b**) PCA, (**c**) traditional GMM and (**d**) proposed GMM method.

	Cointegration	PCA	GMM	Proposed GMM method
FPR	1.25%	1.61%	4.54%	0.60%
FNR	98.33%	99.81%	0%	0%

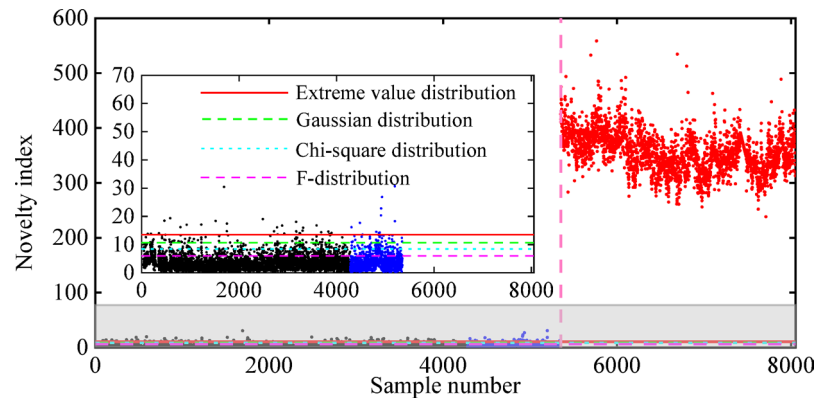**Table 15**. FPR and FNR obtained based on different damage detection methods.

The choice of threshold determination method is crucial for reliable damage detection. Figure 32 compares four threshold methods—Gaussian distribution, chi-square distribution, F distribution, and extreme value distribution—with their respective FPR and FNR summarized in Table 16. The extreme value distribution method demonstrated superior performance, achieving the most accurate results in damage identification. While the other three methods produced slightly lower FNRs, their higher FPRs increased the likelihood of false alarms during the undamaged phase. This limitation arises from the assumption that real-world data strictly follow predefined statistical distributions. The extreme value distribution method, by determining thresholds based on the extreme values of the data, adapts more effectively to the inherent characteristics of the data, reducing false alarms and improving reliability in practical applications.

Limitations and drawbacks
From the above results, we can see that the shortcomings of this method are as follows:

(i) Insensitive to minor damage. For instance, in the example of Fig. 11, the algorithm cannot detect the damage occurrence in damage case 1 with the minor damage (23.5 g). An improved approach is to propose a more sensitive damage indicator to identify the minor damage, such as weighted Mahalanobis distance or threshold-normalized Mahalanobis distance[51].

(ii) The obtained MMSD time series is non-stationary. As shown in Figs. 11 and 23, the minimum Mahalanobis distance obtained by the proposed method has non-stationary fluctuations, which indicates that the environmental effects may not be completely removed. A possible solution to this problem consists in using the switching cointegration method to reduce the nonstationary[22].

Fig. 32. Damage identification results obtained by different threshold calculation methods.

	Extreme value distribution	Gaussian distribution	Chi-square distribution	F distribution
FPR	0.46%	1.50%	4.29%	12.58%
FNR	0%	0%	0%	0%

Table 16. FPR and FNR obtained by different threshold calculation methods.

(iii) The need for sufficient training data under varying environments. As shown in Fig. 27, inaccurate damage detection is obtained when using 40% and 60% of the natural frequencies as training data, while damage can be reliably detected when using 80% and 100% of the natural frequencies as training data. This implies that insufficient training data will lead to the failure of the proposed damage identification method.

Conclusion

In this article, a novel method for detecting structural damage under changing environments is proposed. The method combines sequential iteration and GMM to achieve accurate and efficient results. Firstly, the initial model parameters and the number of Gaussian components in the GMM are determined using the sequential iteration approach. This step helps in setting up the foundation for our analysis. Next, the EM algorithm is employed to construct a baseline GMM based on the natural frequencies observed under varying environmental conditions. This baseline GMM serves as a reference for identifying deviations caused by structural damage. To quantify the extent of novelty or deviation from the baseline, we calculate the MSD between each sample point and the center of each Gaussian component in the GMM. The minimum MSD is then utilized as the novelty index, enabling us to detect and pinpoint structural damage effectively. To validate the practicality and effectiveness of the proposed GMM damage detection method, the frequency measurements from two real-world structures (i.e. the wooden truss bridge and the Z24 bridge) are employed.

According to the results, the detailed conclusions can be drawn as follows: (1) The proposed method effectively discriminates between damaged and undamaged states in real-world bridges with severe environmental variability. It exhibits high detectability and few false alarms, even when the observations are affected by high measurement noise. (2) The proposed GMM damage detection method is not sensitive to the initial point selection. It consistently yields stable and reliable detection results, even when different initial points are used in the sequential iteration step. (3) Comparative studies of different methods demonstrate that the proposed approach outperforms conventional cointegration and PCA methods. It shows a smaller FPR and FNR while achieving higher damage detectability. (4) Regarding the issue of learning sample size, the comparative study on the Z24 Bridge reveals that more reliable and appropriate damage detection results are obtained when the training dataset covers a wider range of environmental conditions. Specifically, the proposed method does not yield reasonable damage detection results when the percentage of training samples is smaller than 60%.

Data availability

All data generated or analysed during this study are included in this published article.

References

1. Han, S. et al. Non-destructive testing and structural health monitoring technologies for carbon fiber reinforced polymers: a review. *Nondestructive Test. Evaluation.* **39**, 725–761 (2024).
2. Lai, X. et al. Digital twin-based non-destructive testing for structural health monitoring of bridges. *Nondestructive Test. Evaluation.* **39**, 57–74 (2024).

3. Pallarés, F. J., Betti, M., Bartoli, G. & Pallarés, L. Structural health monitoring (SHM) and nondestructive testing (NDT) of slender masonry structures: A practical review. *Constr. Build. Mater.* **297**, 123768 (2021).
4. Hou, R. & Xia, Y. Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. *J. Sound Vib.* **491**, 115741 (2021).
5. Ceravolo, R., Coletta, G., Miraglia, G. & Palma, F. Statistical correlation between environmental time series and data from long-term monitoring of buildings. *Mech. Syst. Signal Process.* **152**, 107460 (2021).
6. Anastasopoulos, D. et al. Vibration-based monitoring of an FRP footbridge with embedded fiber-Bragg gratings: influence of temperature vs. damage. *Compos. Struct.* **287**, 115295 (2022).
7. Xia, Y., Hao, H., Zanardo, G. & Deeks, A. Long term vibration monitoring of an RC slab: temperature and humidity effect. *Eng. Struct.* **28**, 441–452 (2006).
8. Li, H., Li, S., Ou, J. & Li, H. Modal identification of bridges under varying environmental conditions: temperature and wind effects. *Struct. Control Health Monit.* **17**, 495–512 (2010).
9. Zhang, Y., Kurata, M. & Lynch, J. P. Long-term modal analysis of wireless structural monitoring data from a suspension Bridge under varying environmental and operational conditions: system design and automated modal analysis. *J. Eng. Mech.* **143**, 04016124 (2017).
10. Aravanis, T. C., Sakellariou, J. & Fassois, S. Damage precise localization under varying operating conditions via the vibration-data-based functional model method: formulation and experimental validation. *J. Sound Vib.* **535**, 117078 (2022).
11. Maes, K., Van Meerbeeck, L., Reynders, E. P. B. & Lombaert, G. Validation of vibration-based structural health monitoring on retrofitted railway Bridge KW51. *Mech. Syst. Signal Process.* **165**, 108380 (2022).
12. Peeters, B. & De Roeck, G. One-year monitoring of the Z24-Bridge: environmental effects versus damage events. *Earthq. Eng. Struct. Dynamics.* **30**, 149–171 (2001).
13. Cross, E., Koo, K., Brownjohn, J. & Worden, K. Long-term monitoring and data analysis of the Tamar Bridge. *Mech. Syst. Signal Process.* **35**, 16–34 (2013).
14. Hu, W. H. et al. Comparison of different statistical approaches for removing environmental/operational effects for massive data continuously collected from footbridges. *Struct. Control Health Monit.* **24**, e1955 (2017).
15. Moser, P. & Moaveni, B. Environmental effects on the identified natural frequencies of the Dowling hall footbridge. *Mech. Syst. Signal Process.* **25**, 2336–2357 (2011).
16. Oh, C. K. & Sohn, H. Damage diagnosis under environmental and operational variations using unsupervised support vector machine. *J. Sound Vib.* **325**, 224–239 (2009).
17. Bayane, I., Leander, J. & Karoumi, R. An unsupervised machine learning approach for real-time damage detection in bridges. *Eng. Struct.* **308**, 117971 (2024).
18. Han, Q., Ma, Q., Dang, D. & Xu, J. Modal Parameters Prediction and Damage Detection of Space Grid Structure under Environmental Effects Using Stacked Ensemble Learning. *Structural Control and Health Monitoring* 5687265 (2023). (2023).
19. Mousavi, M. & Gandomi, A. H. Structural health monitoring under environmental and operational variations using MCD prediction error. *J. Sound Vib.* **512**, 116370 (2021).
20. Wah, W. S. L., Chen, Y. T. & Owen, J. S. A regression-based damage detection method for structures subjected to changing environmental and operational conditions. *Eng. Struct.* **228**, 111462 (2021).
21. Soo, L., Wah, W. & Xia, Y. Elimination of outlier measurements for damage detection of structures under changing environmental conditions. *Struct. Health Monit.* **21**, 320–338 (2022).
22. Huang, J. Z., Li, D. S. & Li, H. N. A new regime-switching cointegration method for structural health monitoring under changing environmental and operational conditions. *Measurement* **212**, 112682 (2023).
23. Worden, K., Baldacchino, T., Rowson, J. & Cross, E. J. Some recent developments in SHM based on nonstationary time series analysis. *Proc. IEEE.* **104**, 1589–1603 (2016).
24. Diao, Y., Sui, Z. & Guo, K. Structural damage identification under variable environmental/operational conditions based on singular spectrum analysis and statistical control chart. *Struct. Control Health Monit.* **28**, e2721 (2021).
25. Wang, X., Li, L., Beck, J. L. & Xia, Y. Sparse bayesian factor analysis for structural damage detection under unknown environmental conditions. *Mech. Syst. Signal Process.* **154**, 107563 (2021).
26. Huang, J., Yuan, S. J., Li, D. & Li, H. -n. A kernel canonical correlation analysis approach for removing environmental and operational variations for structural damage identification. *J. Sound Vib.* **548**, 117516 (2023).
27. Yan, A. M., Kerschen, G., De Boe, P. & Golinval, J. C. Structural damage diagnosis under varying environmental conditions - part II: local PCA for non-linear cases. *Mech. Syst. Signal Process.* **19**, 865–880 (2005).
28. Yan, A. M., Kerschen, G., De Boe, P. & Golinval, J. C. Structural damage diagnosis under varying environmental conditions—part I: a linear analysis. *Mech. Syst. Signal Process.* **19**, 847–864 (2005).
29. Reynders, E., Wursten, G. & De Roeck, G. Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. *Struct. Health Monit.* **13**, 82–93 (2014).
30. Santos, A., Figueiredo, E., Silva, M., Sales, C. & Costa, J. Machine learning algorithms for damage detection: Kernel-based approaches. *J. Sound Vib.* **363**, 584–599 (2016).
31. Jin, S. S. & Jung, H. J. Vibration-based damage detection using online learning algorithm for output-only structural health monitoring. *Struct. Health Monit.* **17**, 727–746 (2018).
32. Sen, D., Erazo, K., Zhang, W., Nagarajaiah, S. & Sun, L. On the effectiveness of principal component analysis for decoupling structural damage and environmental effects in Bridge structures. *J. Sound Vib.* **457**, 280–298 (2019).
33. Bisheh, H. B. & Amiri, G. G. Structural damage detection based on variational mode decomposition and kernel PCA-based support vector machine. *Eng. Struct.* **278**, 115565 (2023).
34. Cross, E., Manson, G., Worden, K. & Pierce, S. Features for damage detection with insensitivity to environmental and operational variations. *Proc. Royal Soc. A: Math. Phys. Eng. Sci.* **468**, 4098–4122 (2012).
35. Shi, H., Worden, K. & Cross, E. J. A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. *Mech. Syst. Signal Process.* **103**, 381–397. https://doi.org/10.1016/j.ymssp.2017.10.013 (2018).
36. Dao, P. B. & Staszewski, W. J. Cointegration and how it works for structural health monitoring. *Measurement* **209**, 112503 (2023).
37. Deraemaeker, A. & Worden, K. A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mech. Syst. Signal Process.* **105**, 1–15 (2018).
38. Figueiredo, E. & Cross, E. Linear approaches to modeling nonlinearities in long-term monitoring of bridges. *J. Civil Struct. Health Monit.* **3**, 187–194 (2013).
39. Figueiredo, E., Radu, L., Worden, K. & Farrar, C. R. A bayesian approach based on a Markov-chain Monte Carlo method for damage detection under unknown sources of variability. *Eng. Struct.* **80**, 1–10 (2014).
40. Hsu, T. Y. & Loh, C. H. Damage detection accommodating nonlinear environmental effects by nonlinear principal component analysis. *Struct. Control Health Monit.* **17**, 338–354 (2010).
41. Abbassi, A., Römgens, N., Tritschel, F. F., Penner, N. & Rolfes, R. Evaluation of machine learning techniques for structural health monitoring using ultrasonic guided waves under varying temperature conditions. *Struct. Health Monit.* **22**, 1308–1325 (2023).
42. Yang, C. & Liu, Y. Detecting the damage of bridges under changing environmental conditions using the characteristics of the nonlinear narrow dimension of damage features. *Mech. Syst. Signal Process.* **159**, 107842 (2021).

43. Mousavi, M. & Gandomi, A. H. Prediction error of Johansen cointegration residuals for structural health monitoring. *Mech. Syst. Signal Process.* **160**, 107847 (2021).
44. Peng, Z., Li, J. & Hao, H. Structural damage detection via phase space based manifold learning under changing environmental and operational conditions. *Eng. Struct.* **263**, 114420 (2022).
45. Sarmadi, H., Entezami, A. & De Michele, C. Probabilistic data self-clustering based on semi-parametric extreme value theory for structural health monitoring. *Mech. Syst. Signal Process.* **187**, 109976 (2023).
46. Sawant, S., Sethi, A., Banerjee, S. & Tallur, S. Unsupervised learning framework for temperature compensated damage identification and localization in ultrasonic guided wave SHM with transfer learning. *Ultrasonics* **130**, 106931 (2023).
47. Du, F., Wu, S., Xing, S., Xu, C. & Su, Z. Temperature compensation to guided wave-based monitoring of bolt loosening using an attention-based multi-task network. *Struct. Health Monit.* **22**, 1893–1910 (2023).
48. Kashyap, P. et al. Unsupervised deep learning framework for temperature-compensated damage assessment using ultrasonic guided waves on edge device. *Sci. Rep.* **14**, 3751 (2024).
49. Sawant, S., Patil, S., Thalapil, J., Banerjee, S. & Tallur, S. Temperature variation compensated damage classification and localisation in ultrasonic guided wave SHM using self-learnt features and Gaussian mixture models. *Smart Mater. Struct.* **31**, 055008 (2022).
50. Kullaa, J. Structural health monitoring under nonlinear environmental or operational influences. *Shock Vib.* **863494** 2014 (2014).
51. Zang, J. G., Huang, H. B., Sun, Z. G. & Wang, D. S. Subdomain principal component analysis for damage detection of structures subjected to changing environments. *Eng. Struct.* **288**, 116265 (2023).
52. Zhang, X., Yang, J., Wang, T., Xie, J. & Tian, L. Quantitative evaluation of crack damage under a variable temperature environment based on a Gaussian mixture model. *Struct. Health Monit.* **22**, 2064–2081 (2022).
53. Ren, Y., Qiu, L., Yuan, S. & Fang, F. Gaussian mixture model and delay-and-sum based 4D imaging of damage in aircraft composite structures under time-varying conditions. *Mech. Syst. Signal Process.* **135**, 106390 (2020).
54. Daneshvar, M. H., Sarmadi, H. & Yuen, K. V. A locally unsupervised hybrid learning method for removing environmental effects under different measurement periods. *Measurement* **208**, 112465 (2023).
55. Ni, Y. Q. & Chen, R. Strain monitoring based Bridge reliability assessment using parametric bayesian mixture model. *Eng. Struct.* **226**, 111406 (2021).
56. Prawin, J. & Bhaskara, G. V. Outlier analysis combined with Gaussian mixture model for structural damage detection. *Materials Today: Proceedings* (2023).
57. Jacobs, W. R., Edwards, H., Li, P., Kadirkamanathan, V. & Mills, A. R. Gas turbine engine condition monitoring using Gaussian mixture and hidden Markov models. *Int. J. Prognostics Health Manage.* **9**, 1–15 (2018).
58. Qiu, L., Fang, F. & Yuan, S. Improved density peak adaptive Gaussian mixture model for damage monitoring in aircraft structures under time-varying conditions. *Mech. Syst. Signal Process.* **126**, 281–304 (2019).
59. Avendaño-Valencia, L. D. & Fassois, S. D. Damage/fault diagnosis in an operating wind turbine under uncertainty via a vibration response Gaussian mixture random coefficient model based framework. *Mech. Syst. Signal Process.* **91**, 326–353 (2017).
60. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
61. Qiu, L., Fang, F., Yuan, S., Boller, C. & Ren, Y. An enhanced dynamic Gaussian mixture model–based damage monitoring method of aircraft structures under environmental and operational conditions. *Struct. Health Monit.* **18**, 524–545 (2019).
62. Kang, Y. *Structural Damage Monitoring Method Based on Adaptive Hierarchical Partitioned Gaussian Mixture Model* (Nanjing University of Aeronautics and Astronautics, 2021).
63. Figueiredo, M. A. T. & Jain, A. K. Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 381–396 (2002).
64. Wang, Q., Ma, S. & Yue, D. Identification of damage in composite structures using Gaussian mixture model-processed lamb waves. *Smart Mater. Struct.* **27**, 045007 (2018).
65. Yi, Z., Pan, N., Liu, Y. & Guo, Y. Study of laser displacement measurement data abnormal correction algorithm. *Eng. Comput.* **34**, 123–133 (2017).
66. Bishop, C. M. & Nasrabadi, N. M. *Pattern Recognition and Machine Learning* Vol. 4 (Springer, 2006).
67. Qiu, T. *EM Algorithm and its Application Research Based on Gaussian Mixture Model* (University of Electronic Science and Technology of China, 2015).
68. Miguel, L. P., Teloli, R. O., da Silva, S. & Chevallier, G. Probabilistic machine learning for detection of tightening torque in bolted joints. *Struct. Health Monit.* **21**, 2136–2151 (2022).
69. Entezami, A., Sarmadi, H. & Behkamal, B. Long-term health monitoring of concrete and steel bridges under large and missing data by unsupervised meta learning. *Eng. Struct.* **279**, 115616 (2023).
70. Sarmadi, H. & Yuen, K. V. Early damage detection by an innovative unsupervised learning method based on kernel null space and peak-over-threshold. *Computer Aided Civil Infrastructure Eng.* **36**, 1150–1167 (2021).
71. Haan, L. & Ferreira, A. *Extreme Value Theory: an Introduction* Vol. 3 (Springer, 2006).
72. Sarmadi, H. & Karamodin, A. A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects. *Mech. Syst. Signal Process.* **140**, 106495 (2020).
73. Kullaa, J. Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mech. Syst. Signal Process.* **25**, 2976–2989 (2011).
74. Anastasopoulos, D., De Roeck, G. & Reynders, E. P. One-year operational modal analysis of a steel Bridge from high-resolution macrostrain monitoring: influence of temperature vs. retrofitting. *Mech. Syst. Signal Process.* **161**, 107951 (2021).

Acknowledgements

Author contributions

Jie-zhong Huang and Jian Yang wrote the main manuscript text; Dong-sheng Li and Wei-chen Kong reviewed the manuscript; Ya-fei Wang collected the experiment data and perform data analysis.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to D.-s.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.