Hindawi
Structural Control and Health Monitoring
Volume 2024, Article ID 5579734, 18 pages
https://doi.org/10.1155/2024/5579734

Research Article

Online Bridge Structural Condition Assessment Based on the Gaussian Process: A Representative Data Selection and Performance Warning Strategy

Yi-Chen Zhu,[1] Yun-Wen Zheng,[1] Wen Xiong,[1] Jiang-Xin Li,[2] C. S. Cai,[1] and Chao Jiang[1]

[1]*Department of Bridge Engineering, School of Transportation, Southeast University, Nanjing 211189, China*
[2]*Nanjing Urban Construction Investment Holding (Group) Co., Ltd, Nanjing 210000, China*

Correspondence should be addressed to Wen Xiong; wxiong12@hotmail.com

Received 5 December 2023; Revised 19 January 2024; Accepted 19 February 2024; Published 7 March 2024

Academic Editor: Songye Zhu

Data-driven methods have now been widely used in structural health monitoring of civil infrastructures thanks to the rapid development of sensor technologies with massive structural and operational condition data. One main issue of data-driven methods is that the computational time increases with the number of monitoring data used, which limits their applications for online structural condition assessment. Focusing on bridge structural health monitoring, this paper proposes a representative data selection strategy for online performance assessment based on Gaussian process models. The proposed method can effectively reduce the required monitoring data size for training, allowing the bridge performance assessment to be conducted in a real-time manner. The method is developed in a probabilistic manner, allowing associated uncertainty of bridge monitoring data to be rigorously considered. A probabilistic warning index is proposed for bridge condition assessment and anomaly detection. The proposed method is validated using synthetic data and applied to structural condition assessment of two full-scale bridges, illustrating the feasibility for real implementations.

1. Introduction

Structural health monitoring (SHM) systems have been widely deployed on civil infrastructures over the past decades thanks to the rapid development of sensor technology [1]. Based on the monitored structural responses and environmental conditions including vibration, stress, temperature, and wind speed, various methods have been developed for condition monitoring and performance assessment [2–5]. Compared to the advancements in field measurements, developing rigorous and reliable SHM methods is still a challenging problem. One main reason is that the structural responses are not only sensitive to structural changes but also to environmental and operational variations (EOVs). This is especially the case for long-span bridges since they are more flexible and prone to vibration under dynamic loading such as wind and traffic with the increase of main span [6, 7]. Correlations between EOVs and structural properties have been modelled using different methods for performance assessment of bridges based on SHM data. Cross et al. [8, 9] proposed a novel approach, namely, cointegration to remove the environmental trends in SHM data. Linear regression models such as linear adaptive filter [10] and Auto Regressivee eXogenous (ARX) models [11] are commonly used to filter out the environmental effect from the long-term structural response data due to the simplicity and efficiency. To better account for the nonlinear behaviour of EOVs, polynomial regression models [12–14] and wavelet-based approaches [15, 16] have also been used for condition monitoring and damage detection of bridges.

Data-driven methods including machine learning, artificial intelligence, and statistical methods have now been gradually accepted as a powerful tool to find out the potential relationship between EOVs and structural responses from

long-term bridge monitoring data. Relationship between bridge modal properties and temperature has been modelled using support vector machine [17] and neural network [18, 19]. Principal component analysis (PCA) has also been used to model the effect of varying load on the natural frequencies of a bridge [20, 21]. A comprehensive review on data mining methods for SHM can be found in [22]. Among others, the Gaussian process (GP) [23] provides promising nonlinear regression tools for modelling the effect of EOVs in SHM which are flexible and fully probabilistic. Applications include modelling dynamic properties of Tamar bridge under different wind speed [24]; detecting sensor fault, structural damage, and EOVs [25]; and modelling both observed and latent EOVs for bridge performance assessment [26].

One main limitation of data-driven methods is that the computational complexity increases significantly with the increase of data, which makes the methods infeasible for online utilization. This is also the case for GP. Training the GP model involves repeated inverse calculation of the covariance matrix, whose computational complexity increases cubically with the number of training data. Approximation methods have been developed to overcome the computational limits. The approximation of the inverse calculation of the covariance matrix [27, 28] and the log marginal likelihood [29] is a common approach. On the other hand, various methods have been developed to approximate the GP model by utilising a smaller training set, which is mainly divided into three categories. The first approach conducts different selection and management criteria to select a subset or active set from the full dataset [30, 31]. The second kind of method involves the use of pseudoinputs as substitutes for the original training data [32, 33]. The third type of method bases on the concept of variational inference and introduces inducing variables [34, 35].

The forgoing methods are generally applicable for GP regression models when the whole training dataset is known. However, in the context of online structure assessment, monitoring data is subjected to stream input into the GP model. Different from the preceding methods, this work focuses on the online GP strategy for bridge structural condition assessment. Several specialised issues should be considered in this context. First, the "whole" training data (if any) are unavailable since the SHM data are collected in a streamed manner. Second, although long-term SHM data can be massive, the redundancy is also significant. In real implementation, most of the SHM data are collected under normal working and environmental conditions without structural anomaly. This provides the opportunity to conduct sparse approximation which uses a subset of data to approximate the full regression model. Third, it will be beneficial to use raw SHM data for training since the one type of SHM data will be used for several different analysis scenarios. Furthermore, structural monitoring data are commonly measured under operation condition where structural responses and EOVs cannot be directly controlled. The resulting uncertainties in the SHM data can be significant and should be well considered when developing structural condition assessment methods.

To address the abovementioned concerns, this work proposes a bridge structural performance assessment method based on GP through a representative monitoring data selection strategy. The proposed method can effectively reduce the size of training data used for GP models, allowing the bridge performance assessment to be conducted in a real-time manner. Compared to other sparse GP methods where induced training data are used, the proposed method keeps the raw SHM data for training. The method is developed in a probabilistic manner, allowing associated uncertainty to be rigorously considered. A probabilistic warning index is proposed for bridge condition assessment and anomaly detection. The proposed method is validated by synthetic data. It also applied the SHM data of two full-scale bridges. Practical issues in real implementations are discussed.

The remaining paper is organised as follows. Section 2 discusses the background of the problem investigated in this work. The framework of proposed representative data selection strategy and online performance warning strategy are introduced in Sections 3 and 4, respectively. The main procedure of the proposed method is then summarised in Section 5. Synthetic data and two full-scale bridge examples are presented in Section 6 to validate the proposed method. The work is concluded in Section 7.

2. Background and Problem Context

The proposed online bridge structural performance assessment method is developed based on GP regression. Its main theory for SHM and computational complexity are first reviewed [23, 36]. To reduce the complexity, the representative data selection strategy is proposed, which will be introduced in the next section.

Assume that the monitored structural response data $\mathbf{y}$ can be represented as

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{x}$ is the environmental and operational variations; $f(\mathbf{x})$ is the structural responses induced by the EOVs with the output denoted as $\mathbf{f}$; and $\boldsymbol{\varepsilon}$ is the modelling error. Consider a kernel-based GP model for describing the functional behaviour between the EOVs and their induced structural response, i.e.,

$$f(\mathbf{x}) \sim \mathrm{GP}(\mathbf{m}(\mathbf{x}, \boldsymbol{\psi}), \mathbf{k}(\mathbf{x}, \boldsymbol{\psi})), \tag{2}$$

where $\boldsymbol{\psi}$ is the associated hyperparameters of the GP model, and $m(.)$ and $k(.,.)$ are the mean and covariance functions (also known as a kernel function) with respect to $\mathbf{x}$ and hyperparameters $\boldsymbol{\psi}$. The outputs of these two functions are denoted as $\mathbf{m}$ and $\mathbf{K}$, respectively.

The probability density function (PDF) of $\mathbf{f}$ given the hyperparameters $\boldsymbol{\psi}$ can be expressed as

$$p(\mathbf{f} \mid \boldsymbol{\psi}) = N(\mathbf{m}, \mathbf{K}). \tag{3}$$

The structure is assumed to be healthy without structural changes at the training stage of the SHM method. In this context, the modelling error is mainly due to sensor noise

and it is sufficient to assume $\boldsymbol{\varepsilon}$ as i.i.d. (independent and identically distributed) Gaussian white noise, whose probability density function can be given by

$$p\left(\boldsymbol{\varepsilon}|\sigma^2\right) = N\left(0, \sigma^2\mathbf{I}\right), \tag{4}$$

where σ^2 is the corresponding variance.

The hyperparameters of the GP model can be inferred by maximising the marginal likelihood function given by

$$\begin{aligned} p(\mathbf{y}\,|\,\mathbf{x}) &= N\left(\mathbf{m}(\mathbf{x}), \mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}\right) \\ &= (2\pi)^{-N/2}\left|\mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}\right|^{-1/2} \exp\left\{-\frac{1}{2}[\mathbf{y} - \mathbf{m}(\mathbf{x})]^T\left[\mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}\right]^{-1}[\mathbf{y} - \mathbf{m}(\mathbf{x})]\right\}. \end{aligned} \tag{5}$$

For prediction, the posterior distribution of test outputs $\mathbf{f}_*$ at test inputs $\mathbf{x}_*$ can be given by

$$p\left(\mathbf{f}_*|\mathbf{y}, \mathbf{x}, \mathbf{x}_*\right) = N\left(\boldsymbol{\mu}_{\mathbf{f}_*}, \mathbf{C}_{\mathbf{f}_*}\right), \tag{6}$$

with

$$\begin{aligned} \boldsymbol{\mu}_{\mathbf{f}_*} &= \mathbf{K}\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}\right]^{-1}\mathbf{y}, \\ \mathbf{C}_{\mathbf{f}_*} &= \mathbf{K}\left(\mathbf{x}_*, \mathbf{x}_*\right) - \mathbf{K}\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}\right]^{-1}\mathbf{K}\left(\mathbf{x}, \mathbf{x}_*\right). \end{aligned} \tag{7}$$

Structural condition assessment can be conducted by investigating the difference between the measured structural responses data and the model predictions. As mentioned before, the modelling error is to be modelled as i.i.d. Gaussian white noise when there is no structural change. When structural changes such as damage and performance degradation occur, additional characteristics will occur in the modelling error, showing different statistical properties. Structural performance warning can then be conducted based on this with the proper threshold set.

As shown in equation (5), inferring the GP model involves inverse calculation of the covariance matrix $\mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2\mathbf{I}$. The computational complexity is of $\mathcal{O}(N^3)$, which grows cubically with the number of training data N. This makes the GP model intractable for online SHM when the dataset is large (which is commonly the case for long-term monitoring data).

3. Representative Data Selection Strategy

In order to conduct SHM based on GP in an online manner, methods should be developed to reduce the computational complexity. As mentioned in the introduction section, SHM data are commonly collected in a streamed manner. When the SHM data are used to train the GP model, the "whole" training data (if any) are unavailable. Approaches based on approximation of the full covariance matrix are not suitable in this case. Considering the redundancy of SHM data, the proposed method in this work is developed based on the sparse framework. The main idea is to select a subset of size m from the whole training dataset of size $n (m < n)$ to approximate the whole GP model. The general framework is given in Algorithm 1.

The algorithm involves two main steps when a new data point comes. First, consider whether the new data point should be added to the current subset. Second, if the new data point is added to the subset, which data point should be removed to keep the size of the subset. The focus here is to design the criteria of these two steps to suit the structural performance assessment based on bridge SHM data.

4. Online Performance Assessment Method

In this section, a bridge structural performance assessment for online implementation is proposed based on the representative data selection framework discussed in the last section. It should be noted that the representative data selection framework only focuses on training. Anomaly detection criteria should also be considered for a performance assessment method. Before considering the training criteria when a new SHM data point comes, the new data point should be first evaluated whether it is collected without structural anomalies. It should be also noted that for some structural anomalies such as degradation or fatigue, such changes are accumulative without mutations. To detect such structural anomalies, another subset of training data is introduced as the reference dataset for performance assessment. The proposed performance assessment strategy should be conducted not only based on the new data point but also the difference between the GP model based on the current data subset and the GP model based on the reference dataset (reference model). Details of the proposed method are given in the following.

First, consider the novelty and anomaly evaluation of a new data point $\{x_{\text{new}}, y_{\text{new}}\}$. The new data are considered to have enough novelty if the difference between itself and the prediction of the current GP model is significant. However, if such difference is excessive, potential structural anomalies should be considered. The current GP model could provide a probabilistic prediction at x_{new}, i.e., $p(f_{\text{new}} \,|\, x_{\text{new}}, \mathbf{D}_m, \psi)$. In order to fully account for the uncertainty in the measured SHM data, the Mahalanobis distance is adopted for the novelty and anomaly evaluation in this work. It is a measure of the distance between a point and a probability distribution, which is well suited in this case. The Mahalanobis distance between $\{x_{\text{new}}, y_{\text{new}}\}$ and $p(y|x_{\text{new}}, \mathbf{D}_m, \psi)$ can be given as

(1) Input: new data point $\{x, y\}$
(2) Persistent: current subset selection $\mathbf{D}_m = \{x_i, y_i\}_{i=1}^m$, hyperparameters of the GP model
(3) Compute the novelty γ of data point $\{x, y\}$
(4) If $\gamma < \gamma_{\text{tol}}$, then
Keep the current subset and GP model
(5) Else
Adding $\{x, y\}$ to $\mathbf{D}_m$ and update the GP model
(6) End if
(7) If data size exceeds m, then
Compute the score (i.e., redundancy, R) of each datapoint $\{x_i, y_i\}$ based on its impact of removal
(8) Remove the data point with the lowest score (i.e., minimum impact of removal, $\text{argmax}(R)$) from the dataset and update the GP model
(9) End if

Algorithm 1: General framework for representative data selection.

$$\gamma = \sqrt{\left(y_{\text{new}} - \mu_{f_{\text{new}}}\right)^{\top} \mathbf{C}_{f_{\text{new}}}^{-1} \left(y_{\text{new}} - \mu_{f_{\text{new}}}\right)}, \tag{8}$$

where $\mu_{f_{\text{new}}}$ and $\mathbf{C}_{f_{\text{new}}}$ are the mean and covariance of $p(f_{\text{new}} \mid x_{\text{new}}, \mathbf{D}_m, \psi)$. For a normal distribution, the value of the Mahalanobis distance is directly related to the confidence interval, which can be used as guidance when setting the threshold for novelty and anomaly evaluation. For example, a Mahalanobis distance of 2 (covering the confidence interval of 95%) can be used as the warning threshold I for performance (denoted as γ_w) while a Mahalanobis distance of 1 can be used as a threshold for subset selection update (denoted as γ_{tol}).

Next, consider the score for impact of removal. In this case, the comparison is between two GP models, which are two probability distributions. Various forms of measures have been developed to quantify the difference between two probability distributions. Among others, Kullback–Leibler divergence is one of the most commonly used measures. As a no upper-bound f-divergence method, KL divergence has advantage in comparing, which introduce as redundancy criteria in this work. The KL divergence of the probability distribution P from Q is defined to be the integral.

$$D_{\text{KL}}(P \| Q) = \int_{-\infty}^{+\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) \mathrm{d}x. \tag{9}$$

In real world, the acquisition data are discrete. Therefore, the GP model is approximated by a k dimensional multivariate normal distribution. For two k dimensional multivariate normal distributions P and Q, with means $\boldsymbol{\mu}$ and $\mathbf{m}$ and with covariance matrices $\boldsymbol{\Sigma}$ and $\mathbf{C}$, the KL divergence between distributions is as follows:

$$D_{\text{KL}}(P \| Q) = \frac{1}{2}\left\{\log\frac{|\mathbf{C}|}{|\boldsymbol{\Sigma}|} + \text{Tr}\left(\mathbf{C}^{-1}\boldsymbol{\Sigma}\right) + (\boldsymbol{\mu} - \mathbf{m})\mathbf{C}^{-1}(\boldsymbol{\mu} - \mathbf{m})^{\top} - k\right\}. \tag{10}$$

In the scenario of accumulative damages, the difference between the reference model (denote as GP_{ref}) and the current model can be used as the warning threshold II. However, the characteristics of asymmetric and unbounded make KL divergence not suitable to measure such distance. In this work, the Hellinger distance is adopted to quantify the discrepancies between two GPs given by

$$H(P, Q) = \sqrt{\frac{1}{2}\int\left(\sqrt{p(P)} - \sqrt{p(\mathrm{Q})}\right)^2 \mathrm{d}x}, \tag{11}$$

where P, Q denote the probability distribution of these two GP models, respectively. For two Gaussian distributions $P \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathrm{Q} \sim N(\mathbf{m}, \mathbf{C})$, the squared Hellinger distance can be written as

$$H^2(P, Q) = 1 - \frac{|\boldsymbol{\Sigma}|^{1/4}|\mathbf{C}|^{1/4}}{|(\boldsymbol{\Sigma} + \mathbf{C})/2|^{1/2}} \exp\left(-\frac{1}{8}(\boldsymbol{\mu} - \mathbf{m})^{\top}\left(\frac{\boldsymbol{\Sigma} + \mathbf{C}}{2}\right)^{-1}(\boldsymbol{\mu} - \mathbf{m})\right). \tag{12}$$

The Hellinger distance considers the complete form of the probability density function and is bounded from 0 (i.e., two distributions are identical) to 1 (i.e., two distributions have no overlay), which is suitable for setting performance warning thresholds. The warning threshold II of the proposed method can then be set based on the Hellinger distance for performance assessment. In real implementation, it is advised to set warning threshold II between 0.6 and 0.8 (e.g., the overlapped area between two normal distributions with the same variance is between 17.6% and 7.3% in this case), which is a tradeoff between modal sensitivity and false alarm rate.

5. Summary of the Procedure

Based on the foregoing discussion, it is feasible to use SHM data for real-time condition assessment of structures. This section summarizes the main steps of the proposed method, which involves two stages. The first stage is the training stage, where the parameters of the proposed model and representative dataset are iteration to obtain from the monitoring data of structure under normal conditions. Meanwhile, two thresholds are computed using statistical methods. The second stage is the monitoring stage, where the structural performance assessment is conducted using the proposed model and test data. Warning index can be calculated from the modelling error and model update will occur simultaneously. The entire process of the proposed method is summarized and illustrated in Figure 1.

5.1. Training Stage

(1) Initialise $\{\psi, \gamma_{\text{tol}}, D_m\}$

(2) Input new monitoring data point $\{x, y\}$,

(3) Calculate γ of $\{x, y\}$ using equation (8),

Case $\gamma < \gamma_{\text{tol}}$: keep the current D_m and the GP model.

Case $\gamma \geq \gamma_{\text{tol}}$: add new data point to D_m and update the GP model.

(4) If the size of D_m changes, calculate R of each data point of D_m using equation (10).

(5) Remove the data point with largest R and update D_m.

(6) Repeat steps (2)–(5) until iterate through the whole dataset.

5.2. Monitoring Stage

(1) $\{\psi, \gamma_{\text{tol}}, \gamma_w, D_m, \text{GP}_{\text{ref}}\}$ from the training stage.

(2) Input new monitoring data point $\{x, y\}$,

(3) Calculate γ of $\{x, y\}$ using equation (8),

Case $\gamma < \gamma_{\text{tol}}$: keep the current D_m and the GP model, then go to step (2).

Case $\gamma_{\text{tol}} \leq \gamma \leq \gamma_w$: add new data point to D_m and update the GP model.

Case $\gamma > \gamma_w$: trigger warnings and go to step (2).

(4) Update D_m and the GP model based on steps (4)-(5) in the training stage.

(5) Calculate $H^2(\text{GP}_{\text{ref}}, \text{GP})$ using equation (12)

(6) If H^2 exceed the warming threshold II, trigger warnings.

(7) Inspect the structure manually to confirm whether a false positive has occurred.

(8) If it is a pseudowarning, update GP_{ref}.

(9) Repeat steps (2)–(8) until iterate through the whole dataset.

6. Illustrative Examples

Three examples are presented in this section to illustrate the proposed method. The proposed method is first validated using synthetic data. SHM data from two full-scale bridges are then considered, illustrating the applicability of the proposed method to real implementations.

In this work, we assume that the mean function of the GP model is zero as no prior mean function is provided and the monitoring data will be normalized before analysis. The covariance $\mathbf{K}$ is assumed to be the squared exponential function, which is commonly used. The entry at position (i, j) is given by

$$K_{i,j} = \sigma_f^2 \exp\left(\frac{\left\|x_i - x_j\right\|^2}{2l^2}\right), \tag{13}$$

where x_i and x_j are the i-th and j-th entry of the observations. σ_f and l are the hyperparameters of the selected covariance function (i.e., $\psi = \{\sigma_f, l\}$).

To validate the applicability of the proposed method in structure assessment, two categories of abnormal scenario are introduced in this section. Abrupt changes in structure response induced by structural cracking or functional failure are first to be simulated. The anomaly simulation data are derived from equation (14) as follows:

$$S' = \frac{1}{1-\gamma} S, \tag{14}$$

where γ is the severity degree of structural, and S' and S are the simulate data and original data, respectively.

It should be mentioned that not all the diseases may cause abrupt changes in structure response; diseases such as structural stiffness degradation may result in accumulate damage to structure. In the second scenario, the incremental response data are simulated by enlarging the original data as follows:

$$S' = \frac{1}{1-\gamma(x-A)/(B-A)} S, \quad x \in [A, B], \tag{15}$$

where $[A, B]$ is the simulation data interval.

6.1. Synthetic Data Example. The key idea of the proposed method is to capture the functional relationship between structural responses and EOVs and assess the state of the structure based on the GP model with reduced training data size. To validate the proposed method, a simulated SHM system is first considered where the functional behaviour between the measured structural response $\mathbf{y}$ and the environmental and operational variations $\mathbf{x}$ are given by

$$\mathbf{y} = 5\mathbf{x}\sin(12\mathbf{x}) + 7\mathbf{x}^3 - \log(\mathbf{x}+1) + \boldsymbol{\varepsilon}. \tag{16}$$

The measured structural responses data are contaminated with measurement noise, which is modelled to be i.i.d. Gaussian distributed with zero mean and standard deviation of 0.2. Figure 2 shows the simulated data and the true function values.

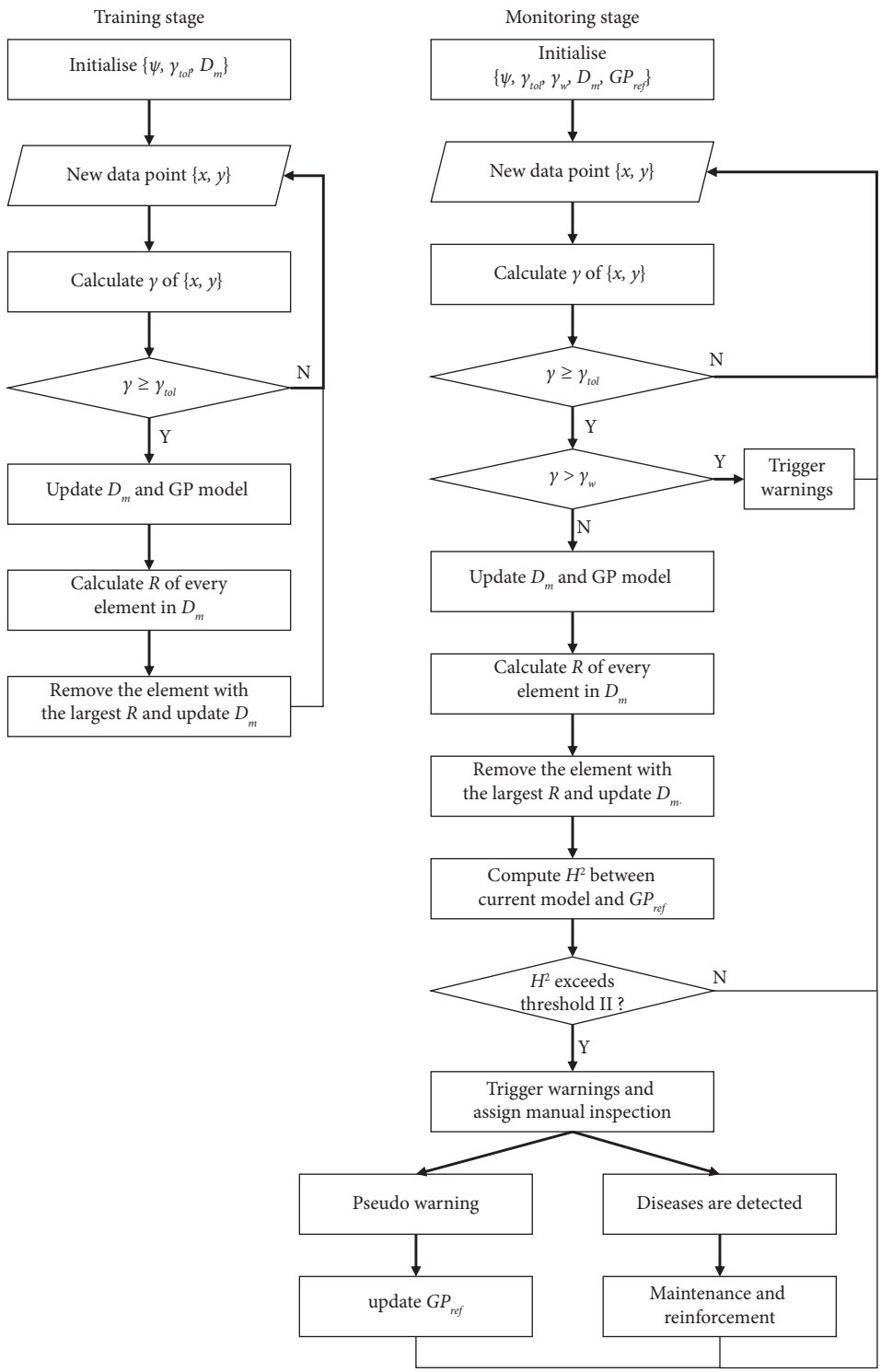

Figure 1: Representative data selection and performance warning strategy process.

To determine the proper size of the representative dataset, simulations are conducted by comparing the predictions of the trained models based on different sizes of training data to the true function values. Figure 3 shows the normalised mean squared error (NMSE) among 30 independent simulations for each training data size condition. It can be seen that with the increase of the training data size, the difference between model predictions and true functional values decreases. When the training data size is larger than 10, further increase of the training data size does not have significant improvement in the model predictions. To balance computational cost and efficiency, the size of the representative dataset is set to be 10 in this example. In real implementation, the dataset size is primarily dictated by the

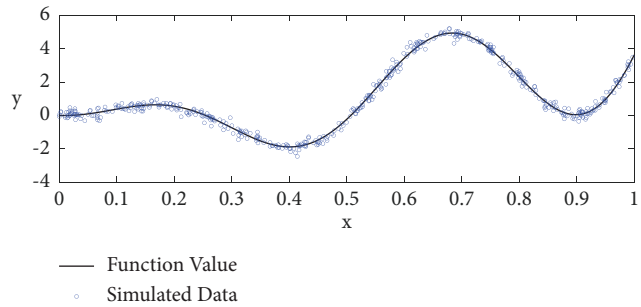

Figure 2: Simulated data with true function values, synthetic data example.

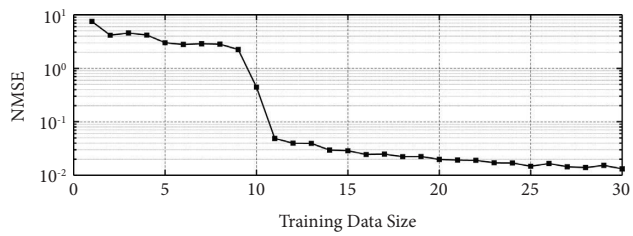

Figure 3: NMSE against the training data size, synthetic data example.

exact storage requirements of bridge data, the time needed for real-time assessment, and the accuracy of the model. While increasing the dataset size can improve model prediction performance, it will also increase the data storage size and evaluation time.

After determining the size of the representative dataset, the updating capability of the proposed method is then investigated. Ten data points are first generated with EOVs within the range of [0, 0.4] to initialise the GP model. Five-hundred data points are then generated within the range of [0, 1] to simulate the streamed SHM data and the proposed strategy is adopted. This simulates the condition where the initial SHM data may not span the whole domain of interest. With the increase of streamed data, more environmental and operation conditions are encountered. Figure 4 shows the model predictions of the initial model and the model with the proposed strategy adopted. Different from the initial training dataset, the updated representative dataset distributed more evenly in the range of interest (i.e., [0, 1]). Although the model predictions based on the initial dataset are more accurate within the range of [0, 0.4], its performance is much worse compared to the model based on the representative dataset in the whole range of interest. Figure 5 shows the comparison between the model based on the whole data and the model based on the representative dataset. The closeness of the model predictions of the two models validates the feasibility of the proposed strategy.

Next, investigate the performance warning capability of the proposed method. The monitoring data are separated into three parts, i.e., the training set, the validation set, and the test set. Figure 6 illustrates the analysis result of the first scenario. The warning threshold I is set to be 1 in this case, covering 68% of the probability under normal condition. In the training stage (i.e., samples 1–200), all the data points are assumed obtained from healthy structures. This means that only the threshold for subset selection update needs to be considered regardless of the judgement of threshold for performance warning. In the monitoring stage (i.e., samples 201–500), the assessment of performance and model update will occur simultaneously.

Compared with minor anomalies, this method is more sensitive to medium and large data anomalies from this figure. In addition, it has a low false-positive rate for training and validation data, which provides well-detecting anomalies capacity.

Figure 7 shows the analysis results of the second scenario with warning threshold II set as 0.6. The tendency of the accumulated offsets can be well captured by the proposed method, which demonstrated its feasibility in this type of anomaly detection. Besides, fluctuations can be found in training and validation intervals. This is mainly because the update of the GP model itself is in constant flux. As long as the threshold is not exceeded, the structures are regarded as healthy.

6.2. Third Nanjing Yangtze River Bridge. Now, investigate the applicability of the proposed method to SHM of full-scale bridges. The Third Nanjing Yangtze River Bridge is first considered. It is a cable-stayed bridge with five spans (main span of 648 m) and an overall length of 1288 m. It is a highway bridge with three lanes in each direction over the Yangtze River in Nanjing, China. Figure 8 shows the front view of the bridge.

The SHM system of these bridges measures more than ten different types of structural responses and environmental and operational conditions (including stress, acceleration, temperature, humidity, and wind speed,) with over two-hundred sensors deployed over the bridge. This example focuses on the stress behaviour of the main girder at the midspan of the bridge under different temperature conditions. Figure 9 shows the sensor locations through the main girder. The SHM data measured in May 2018 are used for analysis. The original sampling rate of stress and temperature is 100 Hz and 1 Hz, respectively, and the measured data are averaged every 10 minutes to suppress the live load effect.

Figure 10 shows the time history of the stress and temperature data. The daily periodic behaviour of stress and temperature indicates the correlativity between them.

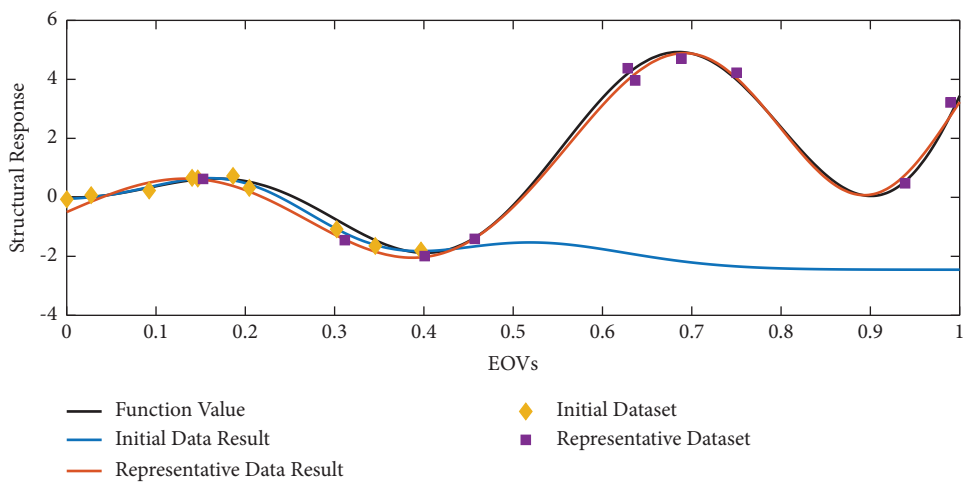

Figure 4: Model predictions based on the initial dataset and the updated representative dataset, synthetic data example.

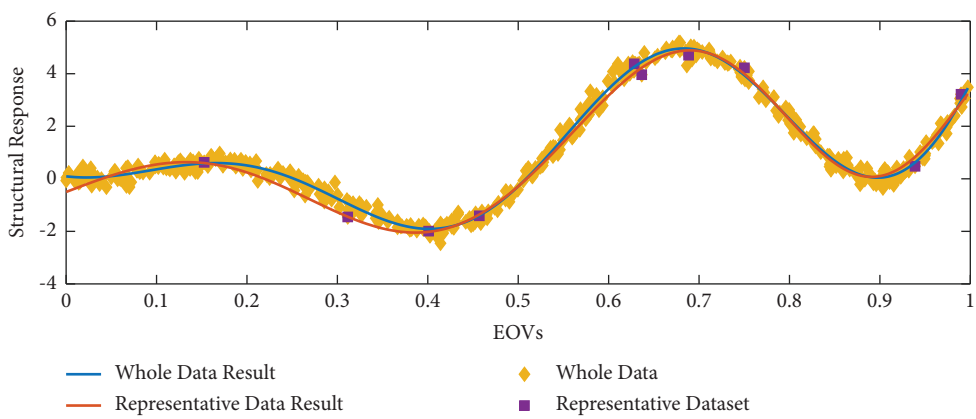

Figure 5: Model predictions based on the whole data and the updated representative dataset, synthetic data example.

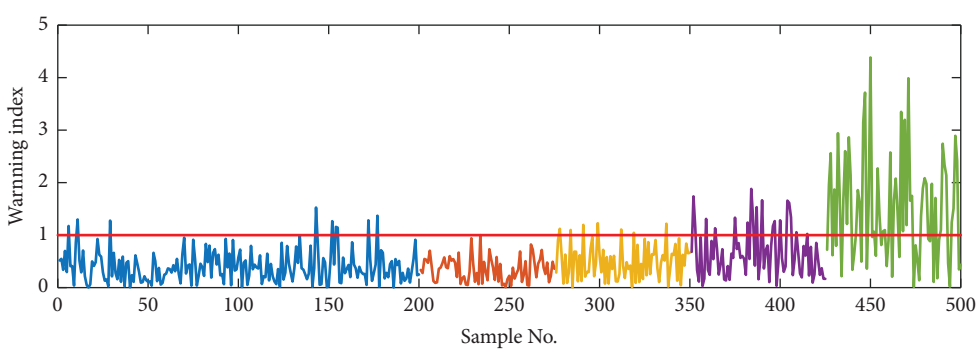

Figure 6: Warning results based on the proposed method with different severity degrees, first scenario, synthetic data example (samples 1–200: training dataset; samples 201–275: validation dataset; samples 276–350: test dataset with 5% severity degree; samples 351–425: test dataset with 10% severity degree; samples 426–500: test dataset with 30% severity degree; red solid line: warning threshold I).

The plots of stress against temperature of different sensors shown in Figure 11 illustrate such correlations. Structural anomaly detection can be conducted if such correlativity can be well described.

It should be noted that although the original data are downsampled, the data amount is still massive (around 4000 data each sensor), which is incapable to conduct online structural performance assessment. Similar to the synthetic data example, investigation is conducted between the model prediction accuracy and training data length. Figure 12 shows the NMSE of model predictions against training data length. It can be seen that 30 data points are sufficient for modelling the relationship between stress and temperature.

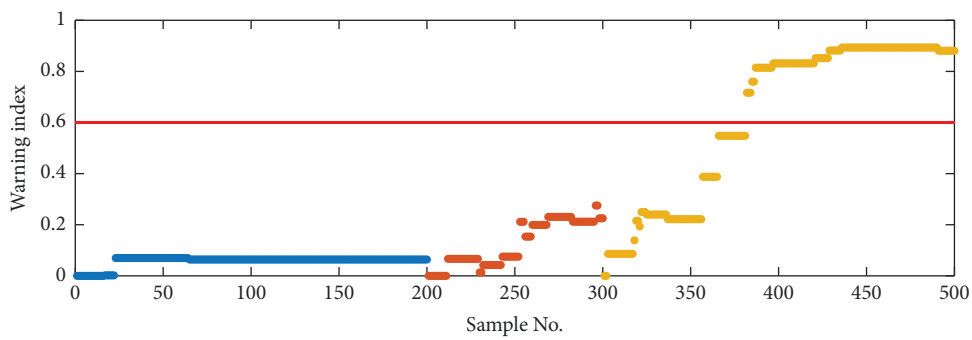

Figure 7: Warning results based on the proposed method with different severity degrees, second scenario, synthetic data example (samples 1–200: training dataset; samples 201–300: validation dataset; samples 301–500: test dataset with severity degree increments from 0 to 30%; red solid line: warning threshold II).

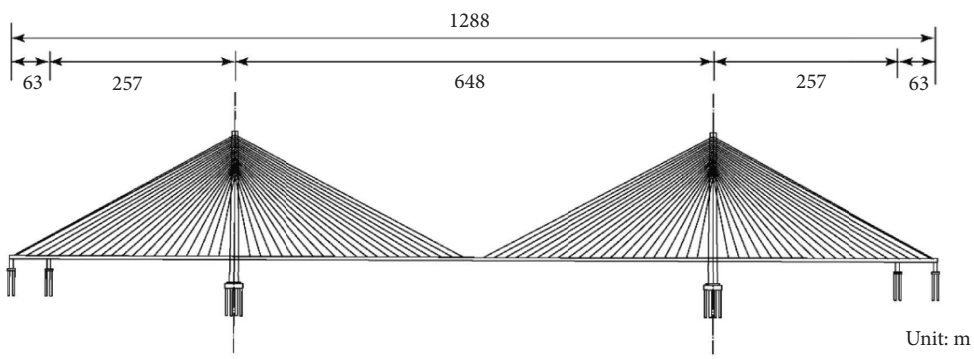

Figure 8: Front view of the Third Nanjing Yangtze River Bridge.

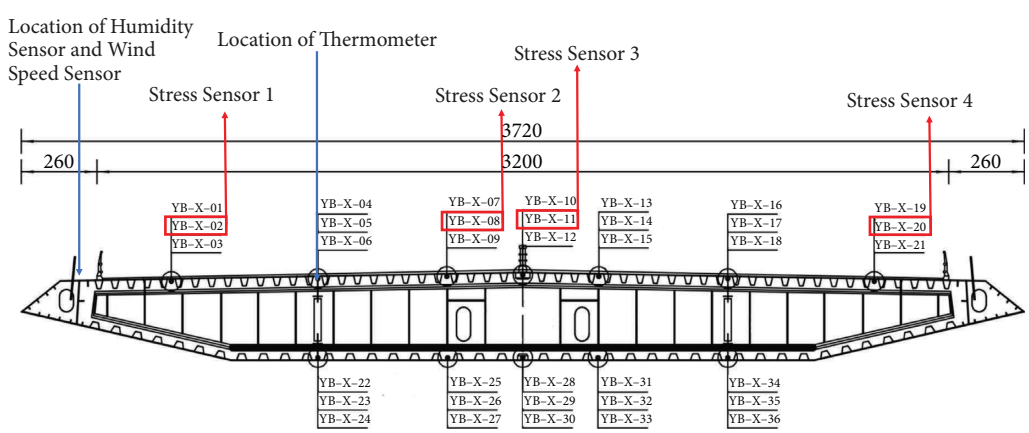

Figure 9: Sensor locations through the main girder, the Third Nanjing Yangtze River Bridge.

Figure 13 shows the model prediction based on the whole dataset and the representative dataset. The NMSE of the model predictions of these two are 0.9855 and 1.050, respectively. The NMSEs between these two datasets are quite close while the computational cost of the proposed method is much smaller, showing the feasibility of the proposed method for real-time SHM in real implementations.

Figure 14 illustrates the results of warning index calculation. In consideration of reducing misjudgement, the warning threshold I is set to be 1.5, which is higher than the synthetic data example. For the 40% severity group, anomalous conditions can be well captured and alerted. However, the sensitivity of anomaly detection at lower severity degrees reduces compared to the synthetic data scenario. One main reason is harsh data acquisition environment on the site which makes noise involved in monitoring data. Different from the i.i.d. Gaussian white noise, measurement noise is chaotic, which increases the difficulty of detecting structural changes.

The analysis results in the second scenario are shown in Figure 15. The warning threshold II of this scenario enlarges to 0.8 for the same reason in the first scenario. The accuracy of prediction is closely related to the difference between the GP model and practical structure. When such difference is

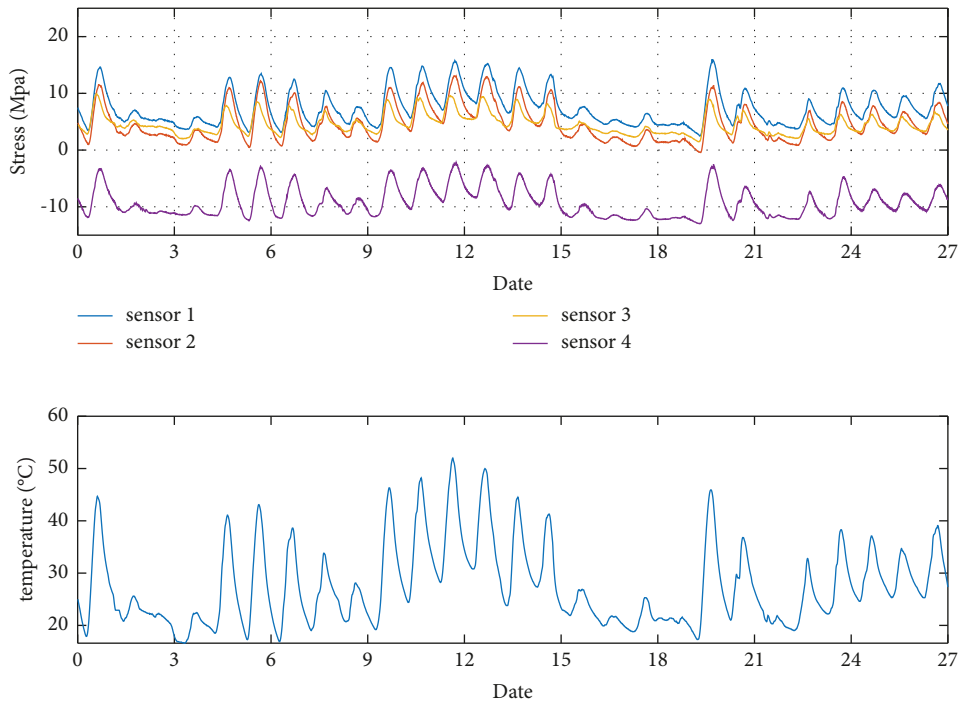

Figure 10: Time history of temperature and stress data of sensors 1–4 used for analysis, the Third Nanjing Yangtze River Bridge.

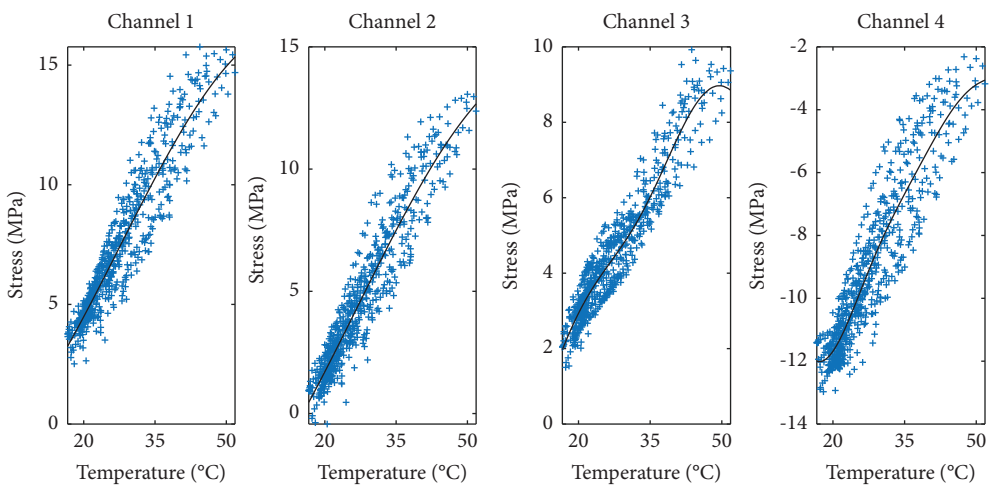

Figure 11: Stress against temperature, the Third Nanjing Yangtze River Bridge (black line: model prediction based on GP).

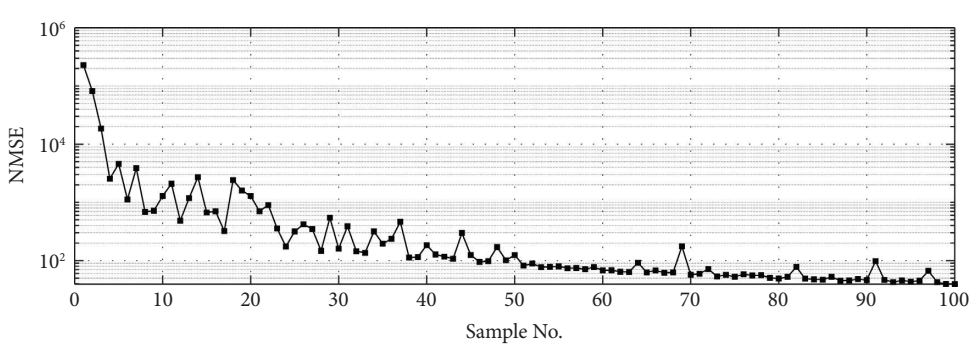

Figure 12: NMSE against the sample number, the Third Nanjing Yangtze River Bridge.

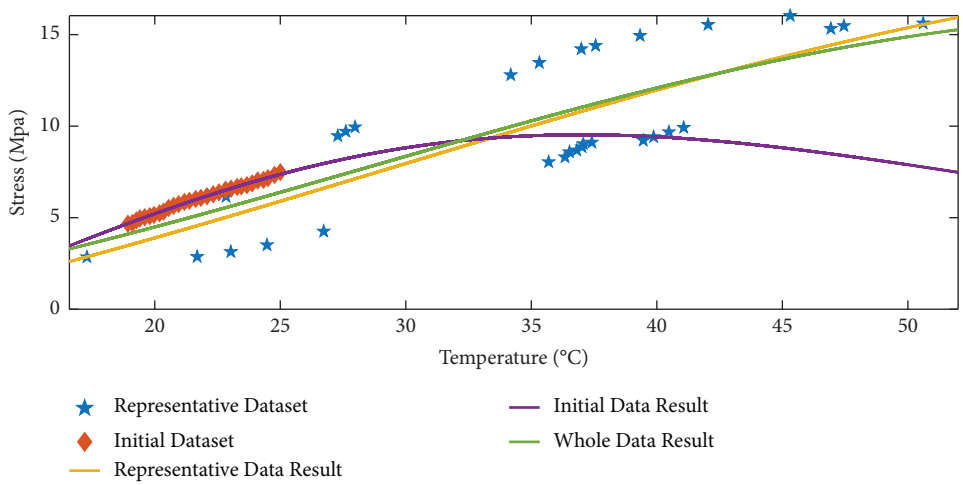

Figure 13: Model predictions based on different training datasets, the Third Nanjing Yangtze River Bridge.

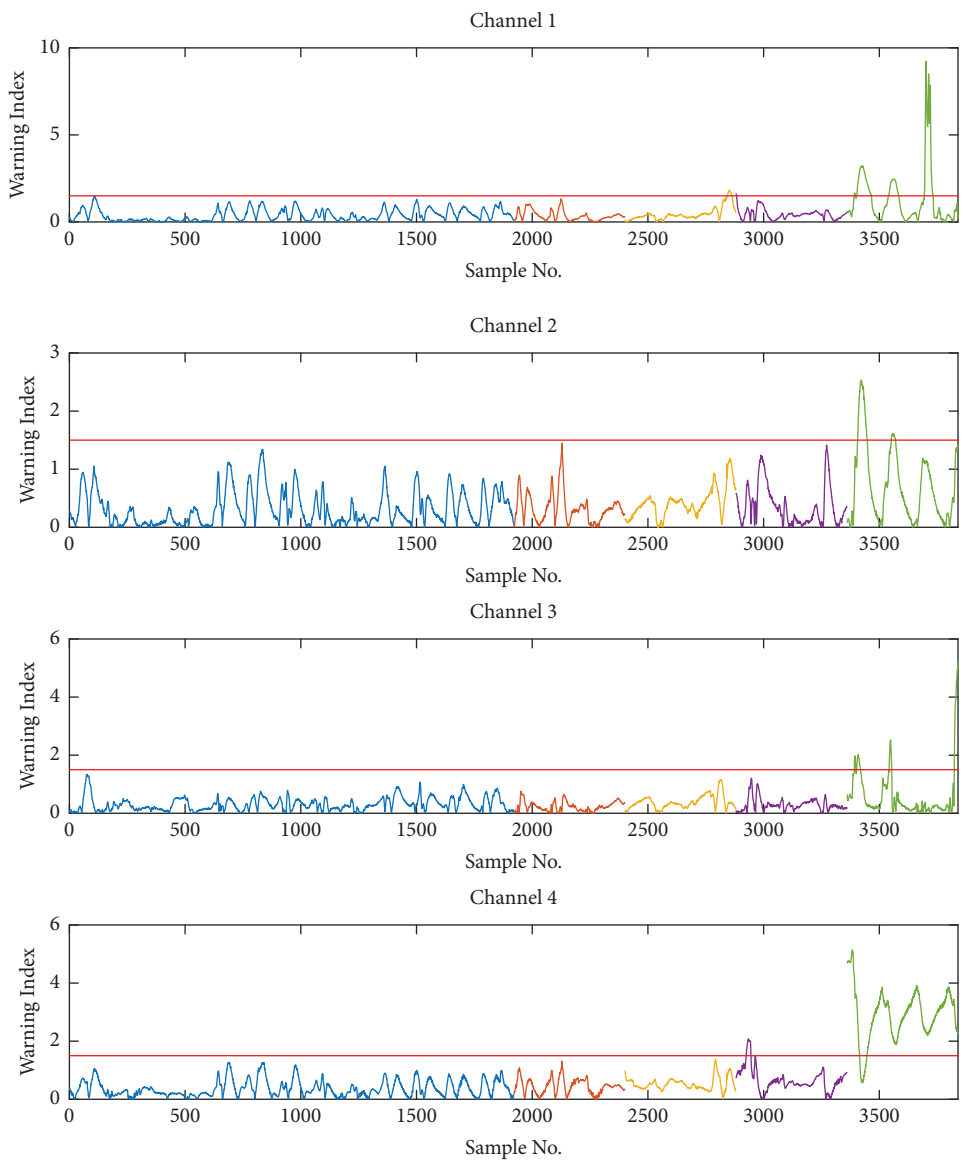

Figure 14: Warning results based on the proposed method with different severity degrees, first scenario, the Third Nanjing Yangtze River Bridge (samples 1–1920: training dataset; samples 1921–2400: validation dataset; samples 2401–2880: test dataset with 10% severity degree; samples 2881–3360: test dataset with 20% severity degree; samples 3361–3840: test dataset with 40% severity degree; red solid line: warning threshold I).

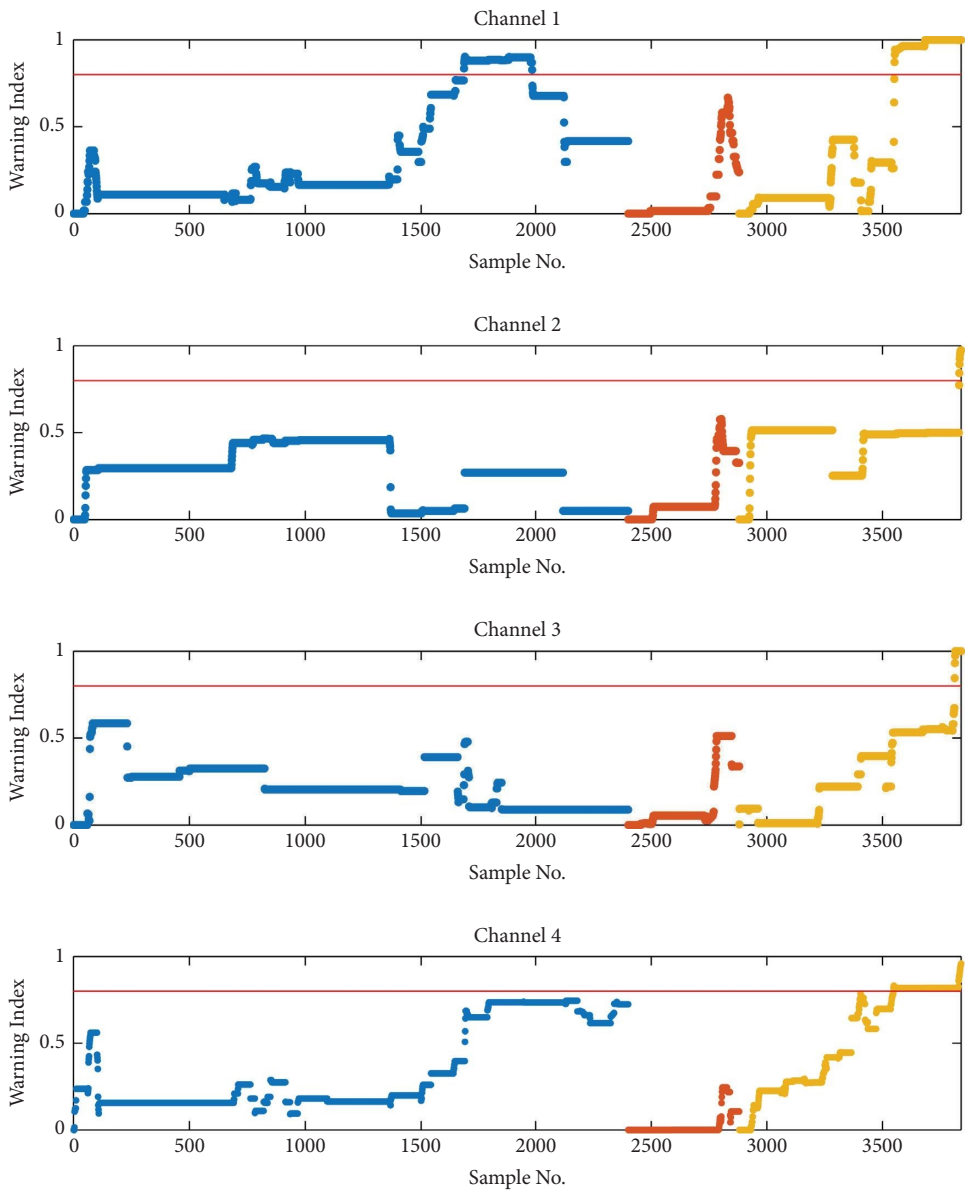

Figure 15: Warning results based on the proposed method with different severity degrees, second scenario, the Third Nanjing Yangtze River Bridge (sample 1–2400: training dataset; samples 2401–2880: validation dataset; samples 2881–3840: test dataset with severity degree increments from 0 to 40%; red solid line: warning threshold II).

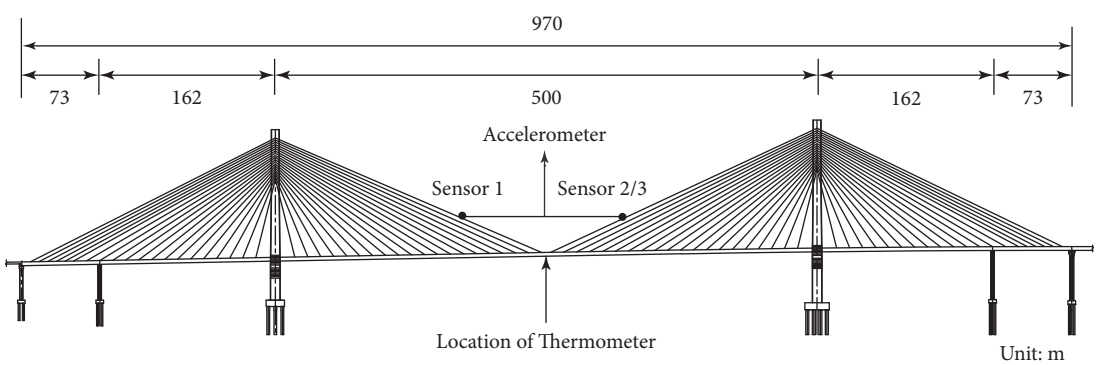

Figure 16: Front view of the Honghe Bridge.

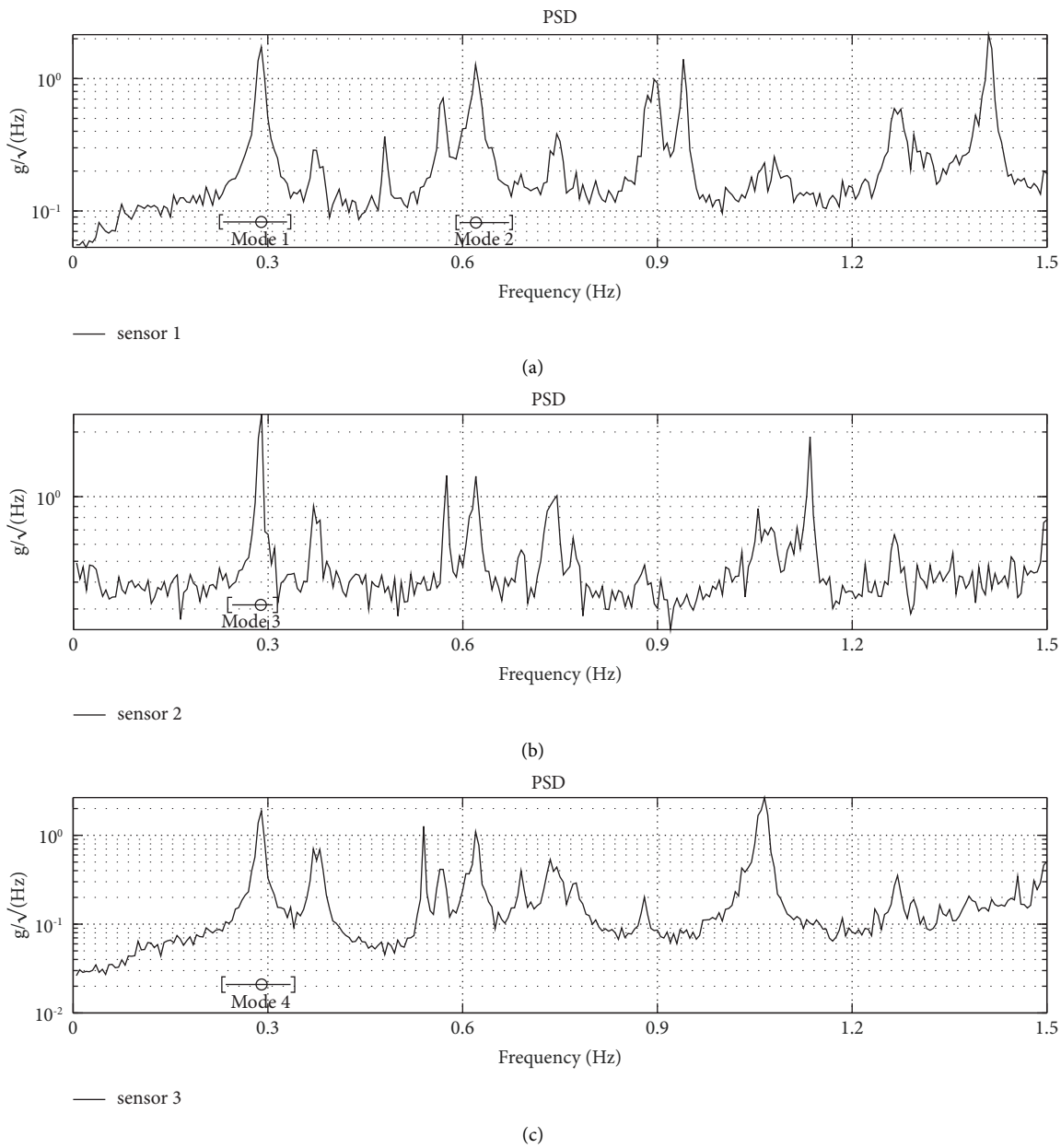

Figure 17: PSD diagram of the measured cable responses, the Honghe Bridge.

excessive, the normal model update may be regarded as deterioration in structure (e.g., the pseudowarning in channel 1). Therefore, the warning triggered by structure damage and model update should be well distinguished in operation. Once the pseudowarning occurred, the reference model needed to update in order to reduce the false positive rate.

6.3. Honghe Bridge, Hongwan Waterway Bridge. The Honghe Bridge is a dual six-lane bridge that spans the Hongwan and Modaomen waterways in Zhuhai. It consists of two double-tower and double-cable plane cable-stayed bridges. The second application is the main channel bridge over the Hongwan waterway has a span arrangement of (73 + 162 + 500 + 162 + 73) m = 970 m. Figure 16 shows the front view of the bridge.

A structural health-monitoring system is in operation to monitor the structural health, including monitoring of environmental loads, structural responses, and traffic. Cables being the crucial structural component of the cable-stayed bridge, its health condition is an essential indicator of that bridge. Diseases such as corrosion will cause a cross-sectional area loss and a reduction of cable load capacity. Therefore, long-term SHM is needed to ensure the cable force within a reasonable range. The vibration-based method is widely used in estimating the cable tension, as the dynamic response can be easily obtained using accelerometers. Besides, cable vibration characteristics are closely linked to their abnormal states. In this example, more than 1200 hours monitoring data from 3 acceleration sensors have been investigated, the locations of which are marked in Figure 16.

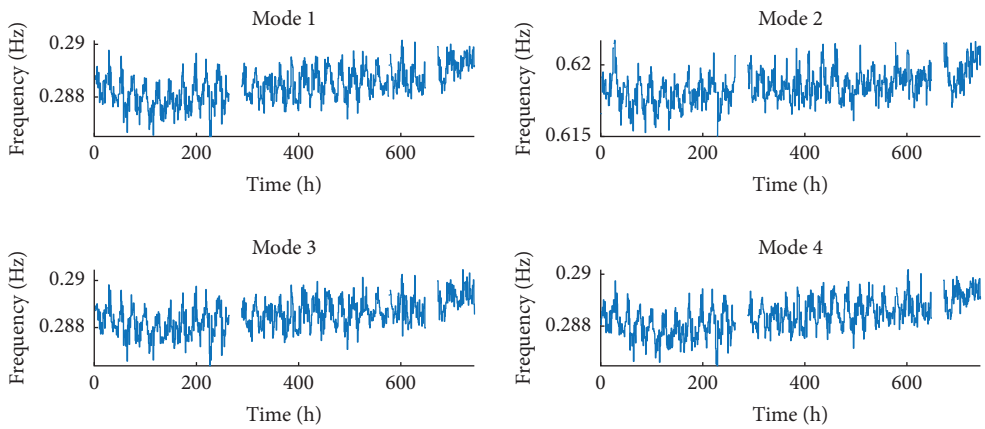

Figure 18: Time history of frequency data used for analysis, the Honghe Bridge.

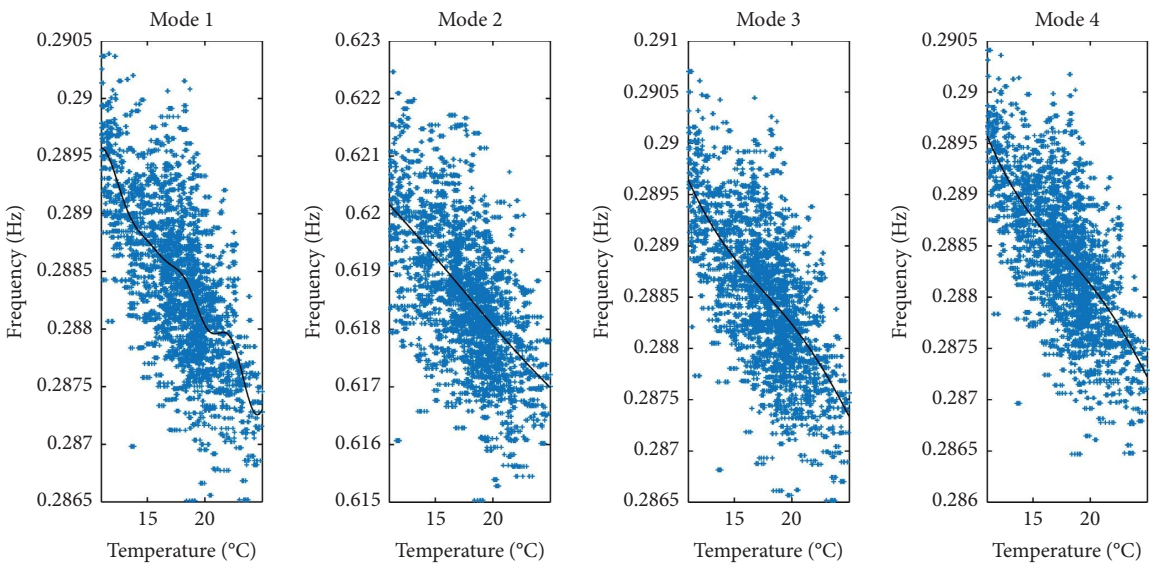

Figure 19: Identified modal frequency against temperature, the Honghe Bridge (black line: model prediction based on GP).

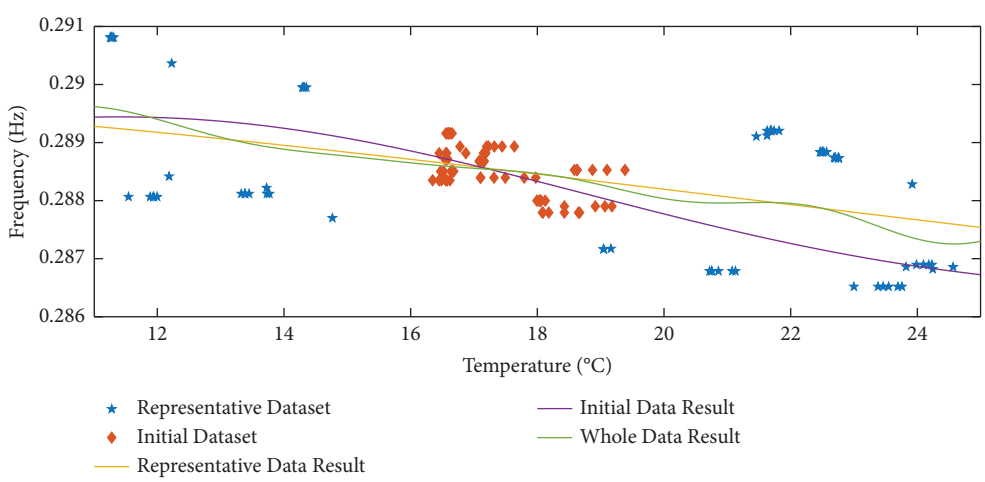

Figure 20: Model predictions based on different training datasets, the Honghe Bridge.

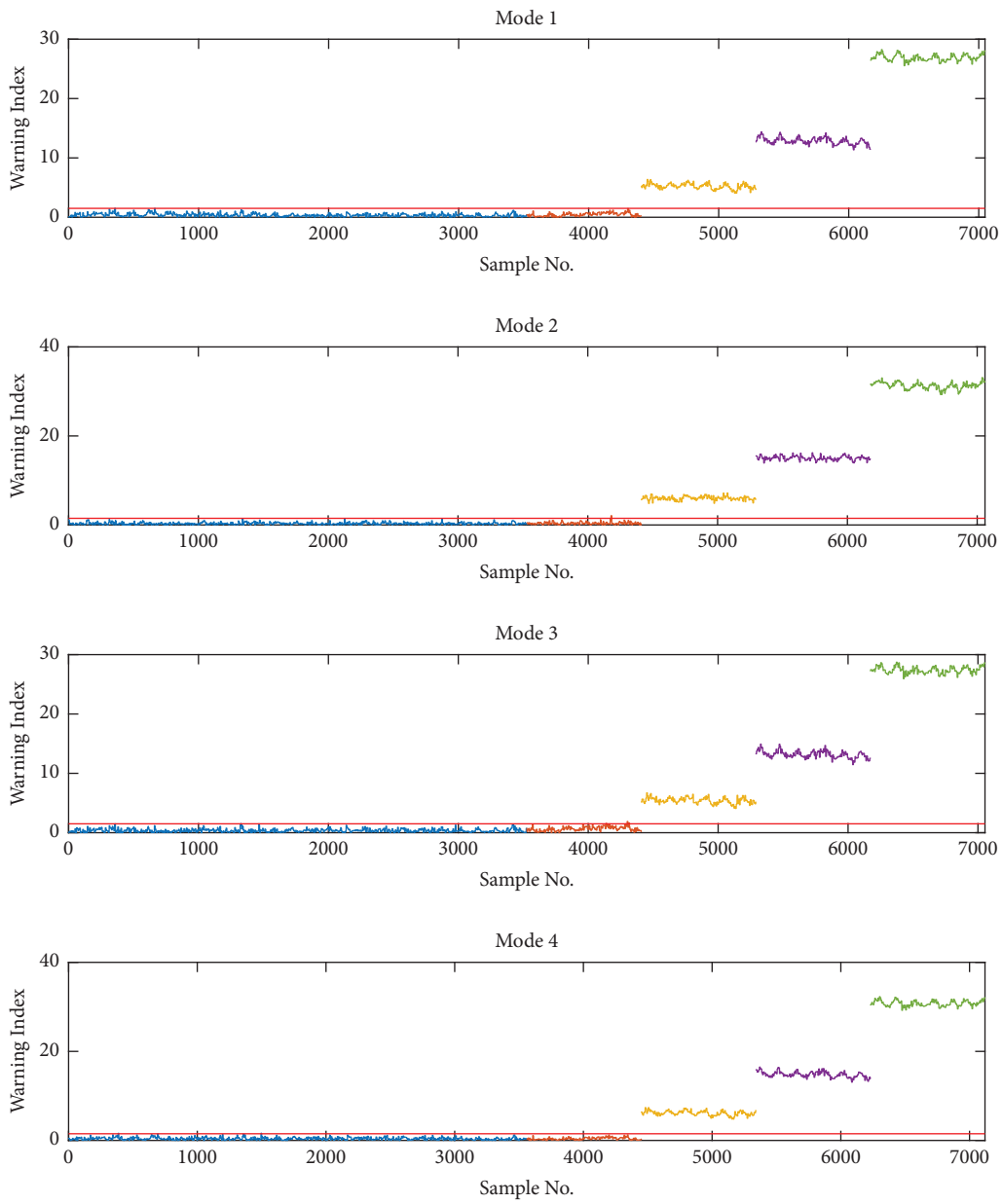

Figure 21: Warning results based on the proposed method with different severity degrees. First scenario, the Honghe Bridge (samples 1–3524: training dataset; samples 3525–4405: validation dataset; samples 4406–5286: test dataset with 2% severity degree; samples 5287–6167: test dataset with 5% severity degree; samples 6168–7048: test dataset with 10% severity degree; red solid line: warning threshold I).

The vibration responses of cables and temperature of the main span are measured at a sampling rate of 50 Hz and 1 Hz, respectively.

Natural frequencies of the cables are identified from nonrepeated 10-minute vibration data while temperature data are averaged over every 10 minutes. Figure 17 shows the power spectral density (PSD) of the investigated cables. The cable modes used for analysis are denoted in this figure. Figure 18 illustrates the time history of the identified natural frequency based on interested modes (mode 1–mode 4) annotated in Figure 17. Similar to temperature variations, changes in the natural frequency show a daily periotic trend. Figure 19 demonstrates the functional relationship between natural frequency and temperature, which shows a negative correlation. However, frequency variations are not only induced by temperature but are also affected by vehicle loads and other ambient loads. It can be evidenced by the discrete distribution of frequency data in Figure 19.

Considering the precision and efficiency requirements, the size of the representative dataset was defined as 70 in this application. Model predictions based on the whole dataset and representative data are illustrated in Figure 20. Compared to the whole dataset, the predictions based on the representative data show a smoother characteristics and similar precision in whole domains. In this application, the data size was reduced from 7048 to 70, which make computation time affordable.

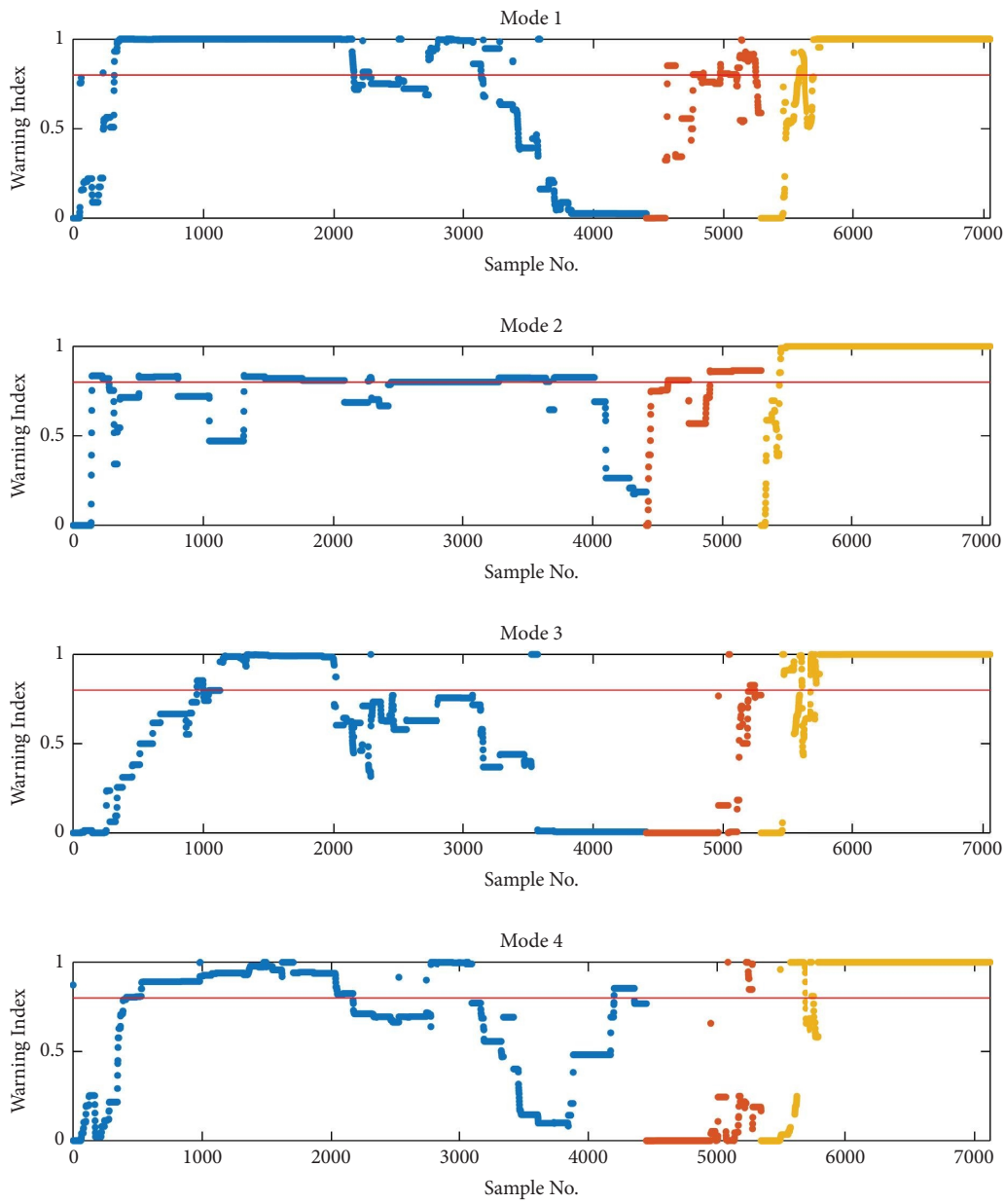

Figure 22: Warning results based on the proposed method with different severity degrees, second scenario, the Honghe Bridge (samples 1–4405: training dataset; samples 4406–5286: validation dataset; samples 5287–7048: test dataset with severity degree increments from 0 to 5%; red solid line: warning threshold II).

Figure 21 presents the analysis results for the first scenario with 1.5 of warning threshold I. It can be seen that the warning threshold I is sensitive to both light and significant changes in structure response. Different from the first case, the warning index in this case has better behaviour and higher precision. This is mainly because the variations of the frequency are less than stress, which make frequency more sensitive to abnormal changes.

Figure 22 presents the analysis results using the proposed method in the second scenario with 0.8 of warning threshold II. Despite fluctuations within the training set, the behaviour of the continuous offset can be well captured. In this case, the proposed method was applied in dynamic analysis in real structure. The nonlinear relationship between temperature and natural frequency can be well described by the proposed method. Compared to the conventional GP model, the proposed method provides similar accuracy with less computation complexity, showing its capability of real-time structural monitoring. In addition, all the warning thresholds mentioned before are not fixed, which are flexible to meet the higher precision or lower false positive rate demand in real implementation.

6.4. Computational Issue. Computational efficiency is one of the most important indicators to the feasibility of the proposed method for real-time SHM and assessment, which will be discussed in this section. All the analyses in

this work are based on the MATLAB 2023a with Intel Core i5 CPU and 48 G RAM. For two full-scale bridges analysis in this work, the computation time of training a GP model based on the whole datasets are 88 s and 561 s, respectively. On the other hand, training a GP model based on the proposed method spends 0.08 s and 0.10 s, respectively. It shows that the proposed method can significantly improve the computational efficiency. Thanks to the condensed size of the training dataset, the proposed method also leads to a lower storage demand than conventional methods.

7. Conclusions

An online structural condition assessment method based on the Gaussian process is proposed in this work for structural health-monitoring data. A representative data selection strategy is proposed specifically for streamed bridge SHM data to reduce the computational time so that the proposed method can be conducted in real-time manner. A performance warning index is also proposed for bridge condition assessment and anomaly detection. The proposed method is developed probabilistically, allowing uncertainties in the monitoring data to be considered.

Three illustrative examples are presented in this paper. The proposed method is first validated using synthetic data. It is shown that the proposed method can significantly reduce the size of training data while preserving sufficient model precision. The structural damage can also be well captured by the proposed method. SHM data measured from two full-scale bridges under varying environmental and operational conditions have also been investigated to illustrate the feasibility of the proposed method in real implementation. Compared to the conventional method which uses the whole training dataset, the proposed method provides similar model prediction accuracy with less computational complexity, allowing the performance assessment to be conducted in a real-time manner. Two different structural anomaly scenarios have been considered in examples. It is shown that the proposed warning index can detect the structural changes. Different thresholds can be set balancing the sensitivity and accuracy of the performance assessment. However, the influence of EOVs is often complex, whereas only one environmental impact (temperature) is taken into account. Future research will focus on estimating structural properties while considering multiple environmental factors and employing more feasible selection criteria.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (Nos. 52208150, 52022021, and 51978160), the Natural Science Foundation of Jiangsu Province (No. BK20220853), and the Key Research and Development Program of Jiangsu Province of China (No. BE2021089). The financial support is gratefully acknowledged.

References

[1] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 365, no. 1851, pp. 303–315, 2007.

[2] P. C. Chang, A. Flatau, and S. C. Liu, "Review paper: health monitoring of civil infrastructure," *Structural Health Monitoring*, vol. 2, no. 3, pp. 257–267, 2003.

[3] J. M. W. Brownjohn, "Structural health monitoring of civil infrastructure," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 365, no. 1851, pp. 589–622, 2007.

[4] H. Sohn, C. R. Farrar, F. Hemez, and J. Czarnecki, "A review of structural health monitoring literature 1996– 2001," in *Proceedings of the Third World Conference on Structural Control*, pp. 1–7, Como, Italy, April 2002.

[5] O. S. Salawu, "Detection of structural damage through changes in frequency: a review," *Engineering Structures*, vol. 19, no. 9, pp. 718–723, 1997.

[6] Y. Fujino and D. Siringoringo, "Vibration mechanisms and controls of long-span bridges: a review," *Structural Engineering International*, vol. 23, no. 3, pp. 248–268, 2013.

[7] S. R. Chen and C. S. Cai, "Accident assessment of vehicles on long-span bridges in windy environments," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 92, no. 12, pp. 991–1024, 2004.

[8] E. J. Cross, K. Worden, and Q. Chen, "Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data," *Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, 2011.

[9] H. Shi, K. Worden, and E. J. Cross, "A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 103, pp. 381–397, 2018.

[10] H. Sohn, M. Dzwonczyk, E. G. Straser, A. S. Kiremidjian, K. H. Law, and T. Meng, "An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge," *Earthquake Engineering & Structural Dynamics*, vol. 28, no. 8, pp. 879–897, 1999.

[11] B. Peeters, J. Maeck, and G. D. Roeck, "Vibration-based damage detection in civil engineering: excitation sources and temperature effects," *Smart Materials and Structures*, vol. 10, no. 3, pp. 518–527, 2001.

[12] Y. L. Ding and A. Q. Li, "Temperature-induced variations of measured modal frequencies of steel box girder for a long-span suspension bridge," *International Journal of Steel Structures*, vol. 11, no. 2, pp. 145–155, 2011.

[13] B. Moaveni and I. Behmanesh, "Effects of changing ambient temperature on finite element model updating of the Dowling Hall Footbridge," *Engineering Structures*, vol. 43, pp. 58–68, 2012.

[14] P. Moser and B. Moaveni, "Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge," *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2336–2357, 2011.

[15] Z. Yu, H. Xia, J. M. Goicolea, and C. Xia, "Bridge damage identification from moving load induced deflection based on wavelet transform and lipschitz exponent," *International Journal of Structural Stability and Dynamics*, vol. 16, no. 05, Article ID 1550003, 2016.

[16] Z. Fan, X. Feng, and J. Zhou, "A novel transmissibility concept based on wavelet transform for structural damage detection," *Smart Structures and Systems*, vol. 12, no. 3_4, pp. 291–308, 2013.

[17] Y. Q. Ni, X. G. Hua, K. Q. Fan, and J. M. Ko, "Correlating modal properties with temperature using long-term monitoring data and support vector machine technique," *Engineering Structures*, vol. 27, no. 12, pp. 1762–1773, 2005.

[18] Y. Q. Ni, H. F. Zhou, and J. M. Ko, "Generalization capability of neural network models for temperature-frequency correlation using monitoring data," *Journal of Structural Engineering*, vol. 135, no. 10, pp. 1290–1300, 2009.

[19] Y. Ding, Y. Deng, and A. Li, "Study on correlations of modal frequencies and environmental factors for a suspension bridge based on improved neural networks," *Science China Technological Sciences*, vol. 53, no. 9, pp. 2501–2509, 2010.

[20] A. M. Yan, G. Kerschen, P. De Boe, and J. C. Golinval, "Structural damage diagnosis under varying environmental conditions- Part I: a linear analysis," *Mechanical Systems and Signal Processing*, vol. 19, no. 4, pp. 847–864, 2005.

[21] A. M. Yan, G. Kerschen, P. De Boe, and J. C. Golinval, "Structural damage diagnosis under varying environmental conditions- Part II: local PCA for non-linear cases," *Mechanical Systems and Signal Processing*, vol. 19, no. 4, pp. 865–880, 2005.

[22] M. Gordan, S.-R. Sabbagh-Yazdi, Z. Ismail et al., "State-of-the-art review on advancements of data mining in structural health monitoring," *Measurement*, vol. 193, Article ID 110939, 2022.

[23] C. K. I. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT press, Cambridge, MA, USA, 2006.

[24] K. Worden and E. J. Cross, "On switching response surface models, with applications to the structural health monitoring of bridges," *Mechanical Systems and Signal Processing*, vol. 98, pp. 139–156, 2018.

[25] J. Kullaa, "Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 25, no. 8, pp. 2976–2989, 2011.

[26] Y. C. Zhu, W. Xiong, and X. D. Song, "Structural performance assessment considering both observed and latent environmental and operational conditions: a Gaussian process and probability principal component analysis method," *Structural Health Monitoring*, vol. 21, no. 6, pp. 2531–2546, 2022.

[27] M. Seeger and C. Williams, "Using the Nyström method to speed up kernel machines," *Advances in Neural Information Processing Systems*, 2001.

[28] A. J. Smola and B. Schölkopf, "Sparse greedy matrix approximation for machine learning," in *Proceedings of the Seventeenth International Conference on Machine Learning*, San Francisco, CA, USA, July 2000.

[29] M. W. Seeger, C. K. I. Williams, and N. D. Lawrence, "Fast forward selection to speed up sparse Gaussian process regression," in *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, pp. 254–261, PMLR, Key West, FL, USA, January 2003.

[30] L. Csató and M. Opper, "Sparse on-line Gaussian processes," *Neural Computation*, vol. 14, no. 3, pp. 641–668, 2002.

[31] N. Lawrence, M. Seeger, and R. Herbrich, "Fast sparse Gaussian process methods: the informative vector machine," *Advances in Neural Information Processing Systems*, vol. 15, 2002.

[32] E. Snelson and Z. Ghahramani, "Sparse Gaussian processes using pseudo-inputs," *Advances in Neural Information Processing Systems*, vol. 18, 2005.

[33] E. L. Snelson, *Flexible and Efficient Gaussian Process Models for Machine Learning*, University of London, University College, London, UK, 2008.

[34] M. Titsias, "Variational learning of inducing variables in sparse Gaussian processes," in *Proceedings of the Artificial Intelligence and Statistics, PMLR*, pp. 567–574, Clearwater Beach, FL, USA, April 2009.

[35] J. Hensman, N. Fusi, and N. D. Lawrence, "Gaussian processes for big data," 2013, https://arxiv.org/abs/1309.6835.

[36] C. Williams and C. Rasmussen, "Gaussian processes for regression," *Advances in Neural Information Processing Systems*, vol. 8, 1995.