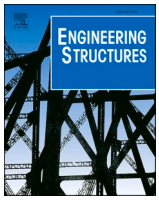
Engineering Structures 284 (2023) 115998

Contents lists available at ScienceDirect

Engineering Structures

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



A data-driven sensor placement strategy for reconstruction of mode shapes by using recurrent Gaussian process regression

Bei-Yang Zhang [a,b], Yi-Qing Ni [a,b,*]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong S.A.R
[b] Hong Kong Branch of the National Engineering Research Center on Rail Transit Electrification and Automation, Hung Hom, Kowloon, Hong Kong S.A.R

ARTICLE INFO

Keywords:
Bridge structure
Data-driven optimal sensor placement
Mode shape reconstruction
Recurrent Gaussian process regression
Greedy algorithm
Cuckoo search algorithm

ABSTRACT

Current Optimal Sensor Placement (OSP) strategies for bridges mostly rely on data from a finite element model rather than from the real structure due to high cost in placing massive sensors for data collection. For large-scale bridges, however, it is difficult to formulate a precise model and thus the OSP strategies building upon a finite element model inevitably suffer from modelling errors. Besides, the finite element model cannot account for real measurement noise. Premised on the fact that it is not expensive to make in-situ trial measurements with a few sensors on a target bridge before deploying a structural health monitoring (SHM) system on it, a data-driven OSP strategy is proposed in this study which aims at accurately reconstructing mode shapes (to facilitate vibration-based structural damage detection) by using only a few vibration sensors to be included in the SHM system. The proposed OSP strategy is also applicable for the upgrade of a long-term SHM system currently deployed on a bridge, by using historical data collected from the current SHM system. To precisely reconstruct mode shapes, a two-stage OSP strategy in terms of Recurrent Gaussian Process Regression (RGPR) is developed, and its performance is validated on a simulation model and a real bridge. In the first stage, the greedy algorithm is leveraged to temporarily deploy sensors on the structure and train accurate RGPR models using the collected data, which are used to afford spatially complete mode shape data for optimization later. Starting from a few sensors temporarily deployed on the bridge, a one-by-one sensor adding procedure is performed to configure increasing sensors until the target is achieved. In the second stage, Cuckoo Search (CS) algorithm is pursued to obtain the globally optimal sensor placement solution, from which the temporarily deployed sensors can be reconfigured to the optimum positions. Both the best sensor quantity and positions are obtained by the proposed OSP strategy.

1. Introduction

Bridges are important civil infrastructure across rivers, valleys, and seas, which are commonly designed with long lifespan and high performance. Due to the traffic loads and environmental effects during the long service life, bridges will inevitably suffer from damage, and the accumulated damage may have a fatal impact to the normal operation of a bridge. To be acquainted with the structural condition in due time, Structural Health Monitoring (SHM) system has been installed on various bridges in the past decades targeting condition monitoring. A typical SHM system generally consists of four subsystems including sensor and data acquisition subsystem, maintenance management subsystem, forewarning and safety evaluation subsystem, and centralized database subsystem [1]. Apparently, the sensor subsystem is the foundation of an SHM system to collect structural responses and environmental effects [2]. Even though this subsystem is crucial, it is impractical to densely deploy sensors on a bridge due to the cost limitation on sensors and hardware and computational limitation in processing massive data [3]. Therefore, the Optimal Sensor Placement (OSP) problem has attracted extensive attention, which aims at deploying a limited number of sensors on proper positions of a bridge to achieve the targets of interest, such as damage detection.

In fact, the OSP problem exists not only in bridges but also in other types of structures, which is essentially a mathematical optimization problem. By building an objective function in alignment with the target of interest, the OSP problem can be transferred into an optimization problem, and then proper optimization algorithms can be performed to solve it. In the previous studies, the target of OSP has been on collecting

* Corresponding author.
E-mail address: ceyqni@polyu.edu.hk (Y.-Q. Ni).

https://doi.org/10.1016/j.engstruct.2023.115998

Available online 22 March 2023

structural information as much as possible with limited measurement points, and a variety of strategies towards this target were proposed [4–18]. For example, Kammer presented the Effective Independence (EI) method for OSP on large space structures [5]. The author proposed to attain a sensor configuration that can collect the largest independent information of the modal coordinates in a structure. Afterwards, the EI method and several improved EI methods were applied to bridge structures for optimal placement of accelerometers [6–10]. The Modal Assurance Criterion (MAC) is a widely used index to identify the orthogonality between the mode shapes of two vibration modes [11], thus it could evaluate the information comprehensiveness on mode shapes perceived from the deployed sensors. The maximum value of the non-diagonal MAC elements was extensively adopted as the optimization target in OSP [12–14]. Besides, the information entropy, which was originated from information theory, was another widely used index in the OSP strategy with a view to reduce the estimation uncertainty of structural parameters [15–18], thereby the configured sensors can help identify the structural properties more accurately. In addition to the purpose of collecting comprehensive structural information, the strategies aiming at damage detection have received great attention as well, since the final target of an SHM system is to identify structural anomalies and to evaluate structural condition [19–22]. The principle of the damage detection-oriented OSP strategy is that the damage information could be included as much as possible in the measured data from the arranged sensors [23–26]. Cobb and Liebst first proposed to employ the concept of damage sensitivity in sensor arrangement strategy [20]. The first order derivatives of structural responses with respect to structural stiffness change were explored as the sensitivity indices [23,27], and sensors were placed at the positions with large sensitivity index values. Some damage-related properties/responses, i.e., modal frequency and mode shapes, and indices such as modal strain energy and flexible curvature [28,29], were also employed to formulate damage sensitive indices for sensor placement. The OSP methods mentioned above were developed mainly for configuring accelerometers, where the modal data was used. Besides accelerometers, other types of sensors such as strain gauges and displacement transducers also need configurations [30,31]. To make the best use of data from different sensors, the OSP strategies for placing multiple types of sensors concurrently became a research hotspot [31–34]. For example, Zhang et al. utilized the relationship between the strain and displacement measurements of a beam structure and fused them for prediction of unobserved structural responses [31]. The minimal prediction error of the structural response was set as the objective function for the arrangement of these two types of sensors. Zhu et al. addressed the problem of placing three types of sensors, i.e., displacement transducers, strain gauges and accelerometers, on a bridge for best reconstruction of structural responses [33].

With the formulated objective function, a proper optimization algorithm is then required to find the sensor configuration with the optimal (maximum/minimum) objective value under a given number of sensors. Therefore, identifying the best sensor configuration and determining the best sensor quantity are two important issues, yet most of the previous studies focused on the first one. The greedy algorithm, or called Forward Sequential Sensor Placement (FSSP) method, is one of the algorithms used to address the OSP problem, which has been applied to numerous practical problems on account of its simple algorithm structure and low calculation burdens [35–38]; moreover, the best sensor quantity can be automatically identified by this algorithm. However, the solution obtained by this algorithm is generally near-optimal rather than optimal [37]. To attain globally optimal solutions of OSP, the heuristic searching algorithms, such as the genetic algorithm, simulated annealing algorithm, particle swarm optimization algorithm and others [12,39–45], were commonly used. The random searching strategy employed in the heuristic algorithms is appealing in addressing such optimization problems, with which the optimal sensor configuration can be obtained with acceptable computational cost. In this method, however, the sensor quantity should be given before the optimization process.

To summarize, in a standard OSP problem, an objective function is first formulated according to the target of interest. Then, the objective function is calculated for each feasible sensor configuration by using the data at the sensor locations. Finally, an optimization algorithm is employed to find the sensor configuration with optimal objective value. Obviously, the structural responses under each feasible sensor configuration are necessary in OSP. In general, such structural responses are obtained from a finite element model. However, using a finite element model inevitably brings about the modelling error problem; meanwhile, the measurement noise, which influences sensor configuration results as well, cannot be properly interpreted from such a model. In a recent study [46], Murugan Jaya et al. aimed to make the expanded mode shapes close to the real mode shapes by considering modelling error and measurement noise. With assumed modelling error and measurement noise, the sensors were deployed at the locations where the observations could help reconstruct the mode shapes accurately. Lin et al. suggested placing sensors at the locations where the vibration signal energy is strong [47]. Such arrangement could improve the signal-to-noise ratio and provide modal parameter estimation with high accuracy. In the above studies, even though the modelling error and measurement noise were assumed, simulation data rather than real measurement data was used [48]. Data-driven methods on addressing the OSP problem have recently attracted attention in various areas such as image recovery, temperature field estimation, mobile sensor navigation [38,49,50]. To the authors' best knowledge, there is a paucity of research on data-driven methods for the OSP problem in structural engineering, partially because of the high cost and low practicability in pre-installing massive sensors for collection of measurement data. In reality, the OSP problem is demanded not only for new bridges but also for the upgrade of SHM systems that have been already deployed on existing bridges for years [51]. How to place sensors on a bridge for trial measurements to facilitate data-driven OSP design of a new SHM system to be deployed on the bridge and how to leverage data collected from an existing SHM system for the upgrade of the system constitute a big challenge.

To address the above, a two-stage data-driven OSP strategy is proposed in this study, in which the target is to accurately reconstruct mode shapes (to facilitate vibration-based damage detection) while requiring only a small number of sensors. The idea is to train a precise surrogate model of the real structure in the first stage using data collected from a few temporarily deployed sensors. Then, by utilizing data generated from the surrogate model, the temporarily deployed sensors can be re-arranged to the optimum positions in the second stage. If the bridge has been instrumented with an SHM system which needs to be upgraded, the required quantity of temporarily deployed sensors can be largely reduced so that the cost of trial measurements is saved. To realize the two-stage OSP strategy, a novel Recurrent Gaussian Process Regression (RGPR) approach is first proposed to build the surrogate model. In this approach, the mode shape data extracted from the time-domain signals is leveraged to enable the estimation of measurement noise, thereby the mode shape components in the spatial domain can be more accurately regressed even if the sensors are sparsely deployed. In line with this approach, the greedy algorithm is pursued in the first stage to instruct the placement of temporary sensors. Embarking on few sensors on the bridge, a one-by-one sensor adding procedure is performed to configure more sensors until precise RGPR models are achieved. These models will be used as surrogate models to the real structure for generating comprehensive response data. In the second stage, the Cuckoo Search (CS) algorithm is executed to find the optimal sensor configuration with data generated by the RGPR models, and the temporarily deployed sensors in the first stage will be re-configured to the optimum positions. Finally, the performance of the proposed RGPR approach is validated on a simulation model and the Ting Kau bridge, and the two-stage OSP strategy is verified on the simulation model.

The rest of this paper is organized as follows. In Section 2, the RGPR approach is proposed, which is followed by validation on a numerical

model and the Ting Kau bridge in Section 3. In Section 4 the data-driven two-stage OSP strategy is presented, and the performance of this strategy is verified on the numerical model in Section 5. Finally, conclusions are drawn in Section 6.

2. Recurrent Gaussian process regression (RGPR)

2.1. Gaussian process regression (GPR) and information theory

2.1.1. GPR

Gaussian Process Regression (GPR) is a Bayesian modelling approach to formulating potentially unknown relationship inherent in data, where no specific function expression needs to be prescribed [52]. The general form of a GPR model is given as:

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

where $f(\boldsymbol{x})$ is an unknown function about the independent variable $\boldsymbol{x}$; ε is a Gaussian white noise term with zero mean $\varepsilon\ N(0, {\sigma_n}^2)$; y is the observation. An assumption made here is that any finite subset of data points $f(\boldsymbol{x})$ corresponding to the different independent variable values fulfils a joint Gaussian distribution. Following this assumption, a prior could be given to the unknown function $f(\boldsymbol{x})$:

$$f(\boldsymbol{x}) \sim\ N(0, K(\boldsymbol{x}, \boldsymbol{x}^{'}))\tag{2}$$

here the mean function is set to $\mathbf{0}$ for simplicity of calculation; $K(\boldsymbol{x}, \boldsymbol{x}^{'})$ is a covariance function, or called kernel function, which describes the potential relations in data points. There are various forms of kernel functions that could be used in different circumstances, in which the square exponential kernel function is most commonly used to represent stationary smooth functions. Given a set of training points $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N\}^T$ in the input space and the corresponding observations $\boldsymbol{y} = \{y_1, \cdots, y_N\}^T$, the likelihood function can be formulated, and then Bayesian inference is performed to obtain the posterior distribution of $f(\boldsymbol{x})$ [52]:

$$p(\boldsymbol{y}|f(\boldsymbol{x}), \boldsymbol{X}, \boldsymbol{\vartheta}) = \prod_{i=1}^{N} N\left(y_i|f(\boldsymbol{x}), \boldsymbol{x}_i, {\sigma_n}^2\right) \tag{3}$$

$$p(f(\boldsymbol{x})|\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{\vartheta}) = \frac{p(\boldsymbol{y}|f(\boldsymbol{x}), \boldsymbol{X}, \boldsymbol{\vartheta})p(f(\boldsymbol{x})|\boldsymbol{X}, \boldsymbol{\vartheta})}{p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\vartheta})} \tag{4}$$

where $\boldsymbol{\vartheta}$ represents a cluster of hyperparameters in $f(\boldsymbol{x})$, which can be optimized by maximizing the marginal likelihood function $p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\vartheta})$ or instead minimizing its negative logarithm form. To make predictions at new positions, the conditional probability can be calculated from the joint distribution between the new positions $\boldsymbol{x}_*$ and training data points $\boldsymbol{X}$:

$$p(f_x, f_*|\boldsymbol{x}_*, \boldsymbol{X}) \sim\ N\left(0, \begin{bmatrix} K(\boldsymbol{X}, \boldsymbol{X}) & K(\boldsymbol{X}, \boldsymbol{x}_*) \\ K(\boldsymbol{x}_*, \boldsymbol{X}) & K(\boldsymbol{x}_*, \boldsymbol{x}_*) \end{bmatrix}\right) \tag{5}$$

$$\begin{aligned} p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, f_x) \sim\ & N\left(K(\boldsymbol{x}_*, \boldsymbol{X})K(\boldsymbol{X}, \boldsymbol{X})^{-1}f_x, K(\boldsymbol{x}_*, \boldsymbol{x}_*)\right. \\ & \left. - K(\boldsymbol{x}_*, \boldsymbol{X})K(\boldsymbol{X}, \boldsymbol{X})^{-1}K(\boldsymbol{X}, \boldsymbol{x}_*)\right) \end{aligned} \tag{6}$$

in which f_x represents $f(\boldsymbol{x})$ for the sake of brevity. Thus, the distribution of predictions becomes

$$p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}, \boldsymbol{\vartheta}) = \int p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, f_x)p(f_x|\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{\vartheta})df_x \sim\ N(f_*|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{7}$$

$$\begin{cases} \boldsymbol{\mu} = K(\boldsymbol{x}_*, \boldsymbol{X})\left(K(\boldsymbol{X}, \boldsymbol{X}) + {\sigma_n}^2 I\right)^{-1}\boldsymbol{y} \\ \boldsymbol{\Sigma} = K(\boldsymbol{x}_*, \boldsymbol{x}_*) - K(\boldsymbol{x}_*, \boldsymbol{X})\left(K(\boldsymbol{X}, \boldsymbol{X}) + {\sigma_n}^2 I\right)^{-1}K(\boldsymbol{X}, \boldsymbol{x}_*) \end{cases}$$

The prediction model $p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}, \boldsymbol{\vartheta})$ is a Gaussian process. Its mean function $\boldsymbol{\mu}$ denotes the best estimation of the data relationship, and the variance function $\boldsymbol{\Sigma}$ provides a confidence interval about the predicted mean function given the training dataset. Owing to the property of providing prediction uncertainty, GPR has been a popular choice in addressing the OSP problem [36–38,50]. This property enables the improvement of modelling accuracy by adding more spatial data without knowing the true modelling errors at these positions. In view of this, GPR is adopted in the present study. With Equation (7), the predicted observation can be expressed with accounting for measurement noise as:

$$p(y_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}, \boldsymbol{\vartheta}) \sim\ N\left(y_*|\boldsymbol{\mu}, \boldsymbol{\Sigma} + {\sigma_n}^2 I_*\right) \tag{8}$$

For the specific problem addressed in this study, the independent variable $\boldsymbol{X}$ represents the spatial coordinates of a bridge. The observations $\boldsymbol{y}$ stand for the modal shape components at different coordinates, and $f(\boldsymbol{x})$ describes the true functional relationship of mode shape. As a consequence, the complete mode shape information of the bridge could be inferred from the prediction model obtained by the GPR method.

2.1.2. Information theory for sensor arrangement

In alignment with Bayesian inference, information theory can help determine the best sensor locations [15,16,37]. The underlying principle is that sensors will be deployed at the locations where the information gain about the target model can be maximized. Thus, the observations collected from the deployed sensors will provide more information for building more accurate Bayesian regression model. A noticeable metric to quantify information is the information entropy, which was originally designed for discrete variables. For a continuous variable, the differential entropy can be used [53]:

$$H_Z = -\int p(z)\ln p(z)dz \tag{9}$$

where $\boldsymbol{Z}$ is a random variable with its probability density function $p(z)$. This index indicates the confidence we have about the true value of the random variable. A large entropy value means low confidence or large uncertainty we know about the true value of this variable. In this study, we integrate information theory with the GPR method. With a set of deployed sensors, a GPR prediction model can be trained by using the collected data, and the uncertainty of predictions at the unsensed positions can be evaluated through the differential entropy. If more sensors are required to train a more precise GPR model, the positions which have large entropy values are preferable to deploy sensors. Thus, through the combination of the differential entropy and GPR method, the OSP problem can be addressed in a data-driven way. In GPR formulation, the predictions at unsensed places are all Gaussian distributions, the differential entropy of the prediction model is a function of the coordinate $\boldsymbol{x}_*$, which can be inferred from Equations (7) and (9) as

$$H(\boldsymbol{x}_*) = -\int p(f_*)\ln p(f_*)df_* = \frac{1}{2}(1 + \ln 2\pi + \ln|\boldsymbol{\Sigma}(\boldsymbol{x}_*)|) \tag{10}$$

2.2. Homoscedastic RGPR

As explored in the previous section, the GPR method is appealing in demonstrating the potential functional relation within dataset, such as the spatial relation of a mode shape in this study. Sufficient spatial training data is necessary for the GPR method to discover the functional relation in data. However, sensors are in general sparsely deployed on a bridge. The spatial observations are probably insufficient for training a precise GPR model. To surmount this problem, a recurrent updating strategy is introduced to the GPR method by leveraging the data in time domain to increase the modelling accuracy. We refer to this approach as Recurrent GPR (RGPR) method.

Assume that a mode shape has a function relation $f(\boldsymbol{x})$ in the spatial domain, and this relation does not change in a time period without the occurrence of structural damage. N sensors are deployed on the structure to collect the mode shape values at different positions, and the spatial observations from D time slots are considered as training data.

$\boldsymbol{y}^d = \{y_1{}^d, \cdots, y_N{}^d\}^T$ represents the observations obtained from N sensors at time d, in which $1 \le d \le D$. By assuming the observation error as a two-dimensional Gaussian white noise in spatial and time domains, the likelihood function at time d can be formulated as follows:

$$p(\boldsymbol{y}^d|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) = \prod_{i=1}^{N} p(y_i{}^d|f_x, \boldsymbol{x}_i, \boldsymbol{\vartheta}) \sim N(f(\boldsymbol{x}), \sigma_n{}^2 I) \tag{11}$$

With the prior assumption about $f(\boldsymbol{x})$, the posterior of $f(\boldsymbol{x})$ given the data $\boldsymbol{y}^1$ collected during the first time slot can be derived by Bayes' theorem as

$$p(f_x|\boldsymbol{X}, \boldsymbol{y}^1, \boldsymbol{\vartheta}) = \frac{p(\boldsymbol{y}^1|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) p(f_x|\boldsymbol{\vartheta})}{p(\boldsymbol{y}^1|\boldsymbol{X}, \boldsymbol{\vartheta})} \tag{12}$$

If data during a new time slot is available, the posterior function obtained in the previous time slot is regarded as the prior and it can be updated. For example, the posterior distribution in time slot $d+1$ can be derived with the likelihood function in time slot $d+1$ and the posterior distribution in time slot d:

$$p(f_x|\boldsymbol{X}, \boldsymbol{y}^{1:(d+1)}, \boldsymbol{\vartheta}) = \frac{p(\boldsymbol{y}^{d+1}|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) p(f_x|\boldsymbol{X}, \boldsymbol{y}^{1:d}, \boldsymbol{\vartheta})}{p(\boldsymbol{y}^{d+1}|\boldsymbol{X}, \boldsymbol{\vartheta})} \tag{13}$$

where $\boldsymbol{y}^{1:(d+1)}$ represents the assembly of observations from time slots 1 to $d+1$. By combining $D-1$ updating steps, the final posterior distribution can be expressed as

$$p(f_x|\boldsymbol{X}, \boldsymbol{y}^{1:D}, \boldsymbol{\vartheta}) = \frac{p(\boldsymbol{y}^{1:D}|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) p(f_x|\boldsymbol{\vartheta})}{p(\boldsymbol{y}^{1:D}|\boldsymbol{X}, \boldsymbol{\vartheta})} \tag{14}$$

where

$$p(\boldsymbol{y}^{1:D}|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) = \prod_{i=1}^{D} p(\boldsymbol{y}^i|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) \prod_{i=1}^{D} N(\boldsymbol{y}^i|f(\boldsymbol{x}), \sigma_n{}^2 I) \tag{15}$$

Similar to the derivation of the prediction model in GPR, the prediction on an unsensed place $\boldsymbol{x}_*$ can be obtained as

$$p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}^{1:D}, \boldsymbol{\vartheta}) \sim N(f_*|\boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) \tag{16}$$

$$\begin{cases} \boldsymbol{\mu}_r = K(\boldsymbol{X}_*, \overline{\boldsymbol{X}}) \left(K(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}}) + \sigma_n{}^2 I\right)^{-1} \boldsymbol{y}^{1:D} \\ \boldsymbol{\Sigma}_r = K(\boldsymbol{X}_*, \boldsymbol{X}_*) - K(\boldsymbol{X}_*, \overline{\boldsymbol{X}}) \left(K(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}}) + \sigma_n{}^2 I\right)^{-1} K(\overline{\boldsymbol{X}}, \boldsymbol{X}_*) \end{cases}$$

where $\overline{\boldsymbol{X}} = \left[\underbrace{\boldsymbol{X}, \boldsymbol{X}, \cdots \boldsymbol{X}}_{D}\right]$. To leverage the information inherent in data acquired from different time slots, the hyperparameter $\boldsymbol{\vartheta}$ is optimized by using all the training data at once instead of step-by-step, which is realized by maximizing the marginal likelihood $p(\boldsymbol{y}^{1:D}|\boldsymbol{X}, \boldsymbol{\vartheta})$. It is worth noting that there is no requirement on the number of samples in the time domain. More samples are recommended for training if the computational burden in the modelling process is accepted.

2.3. Heteroscedastic RGPR

As given in the GPR assumption in Equation (1), the observation is generally regarded as a linear addition of a true model term and a measurement noise term with a constant variance, which implies that the variance of measurement noise keeps invariant everywhere. However, according to the observations made during modal and vibration tests on bridge structures, the measurement noise may have a functional relationship to the modal shape values [54]. Larger mode shape components are likely to be contaminated with higher noise values. Therefore, the heteroscedastic noise should be considered with a view to achieve more accurate modelling results. As such, the assumption of the data model in Equation (1) should be modified as:

$$y = f(\boldsymbol{x})(1+\varepsilon) = f(\boldsymbol{x})(1 + a \cdot \epsilon) = f(\boldsymbol{x}) + af(\boldsymbol{x}) \cdot \epsilon \tag{17}$$

where ϵ is a standard Gaussian distribution with zero mean and unit variance; $af(\boldsymbol{x})$ is the standard deviation of noise which varies with the spatial coordinate $\boldsymbol{x}$; $f(\boldsymbol{x})$ here is termed Noise-Related Function (NRF); a is a hyperparameter. The noise standard deviation is assumed here to be linear to the mode shape value. However, it is found in recent studies that such regularity may not be true [46,54]. For modelling flexibility, a nonlinear relationship can also be applied by changing NRF to nonlinear form in the proposed heteroscedastic RGPR method. Equation (17) implies that the standard deviation of measurement noise varies with the change of spatial coordinate. The prior distribution of the latent function $f(\boldsymbol{x})$ is still assumed as given in Equation (2). With a group of training data $(\boldsymbol{X}, \boldsymbol{y}^d)$ from N sensors in time slot d, the likelihood function can be obtained as

$$p(\boldsymbol{y}^d|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) = \prod_{i=1}^{N} p(y_i^d|f_x, \boldsymbol{x}_i, \boldsymbol{\vartheta}) \sim N(\boldsymbol{y}^d|f(\boldsymbol{x}), a^2 diag(f^2(\boldsymbol{x}))) \tag{18}$$

where $diag(f^2(\boldsymbol{x})) = diag\left(\underbrace{f^2(\boldsymbol{x}), f^2(\boldsymbol{x}), \cdots, f^2(\boldsymbol{x})}_{N}\right)$. Then the posterior distribution can be obtained through Bayesian inference. However, the derivation of posterior distribution is analytically intractable due to the existence of NRF in the noise term. Therefore, an iterative strategy is proposed to use a known function $f_p(\boldsymbol{x})$ to replace $f(\boldsymbol{x})$ in NRF:

$$y = f(\boldsymbol{x})(1 + a \cdot \epsilon) = f(\boldsymbol{x}) + af(\boldsymbol{x}) \cdot \epsilon \approx f(\boldsymbol{x}) + af_p(\boldsymbol{x}) \cdot \epsilon \tag{19}$$

Then the likelihood function in time slot d ($d \in [1, D]$) is

$$p(\boldsymbol{y}^d|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) = \prod_{i=1}^{N} p(y_i^d|f_x, \boldsymbol{x}_i, \boldsymbol{\vartheta}) \sim N\left(f(\boldsymbol{x}), a^2 diag\left(f_p^2(\boldsymbol{x})\right)\right) \tag{20}$$

Thus, the posterior distribution can be derived via Bayesian inference. With observations obtained from D time slots, the posterior distribution of $f(\boldsymbol{x})$ can be elicited as

$$p(f_x|\boldsymbol{X}, \boldsymbol{y}^{1:D}, \boldsymbol{\vartheta}) = \frac{p(\boldsymbol{y}^{1:D}|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) p(f_x|\boldsymbol{\vartheta})}{p(\boldsymbol{y}^{1:D}|\boldsymbol{X}, \boldsymbol{\vartheta})} \tag{21}$$

in which

$$p(\boldsymbol{y}^{1:D}|f_x, \boldsymbol{X}, \boldsymbol{\vartheta}) \sim \prod_{i=1}^{D} N\left(\boldsymbol{y}^i|f(\boldsymbol{x}), a^2 diag\left(f_p^2(\boldsymbol{X})\right)\right) \tag{22}$$

Clearly, if the presumed function $f_p(\boldsymbol{x})$ is in alignment with the predicted mean function $f(\boldsymbol{x})$, i.e., discrepancy between the two functions is very small, the regressed result in Equation (21) can be deemed as the true outcome. We find that the predicted mean function of heteroscedastic RGPR model generally has very tiny discrepancy with $f_p(\boldsymbol{x})$ when the predicted mean function of homoscedastic RGPR model is employed as $f_p(\boldsymbol{x})$. Therefore, the homoscedastic RGPR method will be applied to obtain $f_p(\boldsymbol{x})$ before using the heteroscedastic RGPR method. As a result, the prediction model of heteroscedastic RGPR is expressed as:

$$p(f_*|\boldsymbol{x}_*, \boldsymbol{X}, \boldsymbol{y}^{1:D}, \boldsymbol{\vartheta}) \sim N(f_*|\boldsymbol{\mu}_h, \Sigma_h) \tag{23}$$

$$\begin{cases} \boldsymbol{\mu}_h = K(\boldsymbol{x}_*, \overline{\boldsymbol{X}}) \left(K(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}}) + a^2 diag\left(f_p^2(\overline{\boldsymbol{X}})\right)\right)^{-1} \boldsymbol{y}^{1:D} \\ \Sigma_h = K(\boldsymbol{x}_*, \boldsymbol{x}_*) - K(\boldsymbol{x}_*, \overline{\boldsymbol{X}}) \left(K(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}}) + a^2 diag\left(f_p^2(\overline{\boldsymbol{X}})\right)\right)^{-1} K(\overline{\boldsymbol{X}}, \boldsymbol{x}_*) \end{cases}$$

It should be noted that the above RGPR formulation is working for single output only. For multi-output problem, i.e., modelling multiple mode shapes in this study, an individual RGPR model needs to be formulated for each mode shape, and the relationship among the multiple outputs is not able to be described. Even though the proposed RGPR method is available only for modelling single-output problem, it has the potential to be combined with multiple output Gaussian processes that can transfer the knowledge across related outputs to enhance prediction

accuracy globally. This issue needs further studies in the future.

In fact, many machine learning methods can be employed to train surrogate models, such as polynomial regression, neural network, relevant vector machine (RVM) and so on. In these methods, the Bayesian regression approaches (i.e., RVM) are preferred in addressing the OSP problem since the prediction uncertainty could help instruct the configuration process. Compared with the existing Bayesian regression approaches, the proposed RGPR yields an explicit functional form, which would facilitate practical application (e.g., for modal-based damage detection); and the time domain data can be leveraged to enhance the spatial modelling results, thus the required sensor quantity can be largely reduced to save on costs.

3. Reconstruction of mode shapes by RGPR

3.1. A numerical three-span continuous beam bridge model

To validate the proposed homoscedastic and heteroscedastic RGPR methods, a numerical three-span continuous beam bridge model is first explored. The layout of the bridge and its cross-section properties are shown in Fig. 1. The mass density and elastic modulus of the material are $2.55 \times 10^3\ kg/m^3$ and $3.45 \times 10^4\ MPa$, respectively. To conduct modal analysis, a finite element model is built with 60 equal length segments, in which each segment is represented by 6 degree-of-freedoms (DOFs) Euler-Bernoulli beam element with 1 m length. Under this mesh size, the mode shapes can be portrayed smoothly, and the model can provide enough candidate sensor positions for the subsequent sensor arrangement problem. For GPR and RGPR modelling, the first three vertical mode shapes are of interest. The mode shape components are first normalized by mass and then scaled up by multiplying a constant to avoid pathological calculation in GPR. As an example, 6 accelerometers are deployed on the bridge by experience to obtain the mode shape data, which is shown in Fig. 1(a).

In the simulation study, a Gaussian white noise is added to the mode shape components to simulate the measurement noise with a changing standard deviation at different positions. The noise standard deviation is assumed to be 10 % of the mode shape magnitude. It is worth mentioning that the mode shape values at supports are zero. These values can also be included for GPR modelling.

The mode shapes reconstructed by the GPR method are shown in Fig. 2(a-c), where 6 sensors are considered. To demonstrate that data sufficiency will affect the modelling accuracy of GPR, the results are also obtained by the GPR method when 61 sensors are considered, which are illustrated in Fig. 2(d-f). Results obtained by the homoscedastic and heteroscedastic RGPR methods are depicted in Fig. 2(g-i) and (j-l), respectively.

In Fig. 2, the grey areas denote the 3σ range of the prediction model of observations y_*; the dash lines denote the 3σ range of estimated noise. If the variance of the predicted observations favorably matches the noise variance, it indicates that the uncertainty of the prediction model is all generated by the measurement noise. When the two variances have large discrepancy to each other, it can be inferred that data in spatial domain are insufficient to train a prediction model with low uncertainty. In the results obtained by GPR with 6 sensors, the estimated noise is too small to be depicted. It can be seen that, by the use of GPR, the prediction models trained with data from 61 positions outperform those trained with data from 6 positions. GPR models trained with more spatial data can get more accurate predicted mean function, and measurement noise can also be well estimated. The prediction models obtained by the homoscedastic and heteroscedastic RGPR methods both have good performance with only 6 sensors. They have similar precision to the GPR models trained using sufficient spatial data, and the predicted mean functions agree well with the true values. Meanwhile, the measurement noise estimated by the heteroscedastic RGPR method is closer to the assumption.

To gain an intuitive insight into modelling accuracy for the scaled mode shapes, the maximum Absolute Error Percentage (AEP) is introduced:

$$\mathrm{AEP}_i = |\mu(x_{*i}) - true(x_{*i})|/\max(|true|) \tag{24}$$

$$maximum\ \mathrm{AEP} = \max_{i\in[1,61]}(\mathrm{AEP}_i) \tag{25}$$

in which i represents the node number in the finite element model; $\mu(x_{*i})$ is the predicted mean function value at position x_{*i}; $true(x_{*i})$ is the true mode shape value at position x_{*i}; and $\max(|true|)$ denotes the maximum value of $|true|$ among all the nodes. The maximum AEPs from different methods are depicted in Fig. 3(a), and the estimated standard deviations of measurement noise are shown in Fig. 3(b). For brevity, only the modelling results with 6 sensors are shown.

In Fig. 3, GPR (min) and GPR (max) represent the minimum and maximum values in the modelling results from 30 time slots, respectively. To summarize, the proposed homoscedastic and heteroscedastic RGPR methods can reconstruct the mode shape components with high accuracy, while the GPR models are largely contingent on the quality of training data. In addition, the measurement noise estimated by the heteroscedastic RGPR method is more precise than the homoscedastic RGPR and GPR methods.

3.2. Ting Kau bridge

The second case study uses real-world data collected from the Ting Kau bridge. The Ting Kau bridge is a four-span cable-stayed bridge with two main spans of 448 m and 475 m respectively, and two side spans of 127 m each. More than 230 sensors have been permanently installed on the bridge to collect the structural responses [55]. In this case study, the vertical mode shape is of interest, so the data collected by the accelerometers which are deployed on the bridge deck is used. As shown in Fig. 4, sensors are placed at eight monitoring positions, and some of the acceleration data collected from these sensors are shown, which will be utilized to evaluate the proposed method. In this structure, supports at the two ends of the bridge constrain the vertical displacement, so the mode shape values at these two positions are equal to zero, which can be used as training data.

To identify the mode shape from the structural responses under ambient vibration excitations, the Operational Modal Analysis (OMA) is employed. The data-driven stochastic subspace identification technique is a powerful OMA identification technique [56], which is adopted in the present study. This technique is to formulate a state-space model for the structure system with the collected time-domain acceleration signals. By assuming the observation noise as a zero mean Gaussian white noise, it

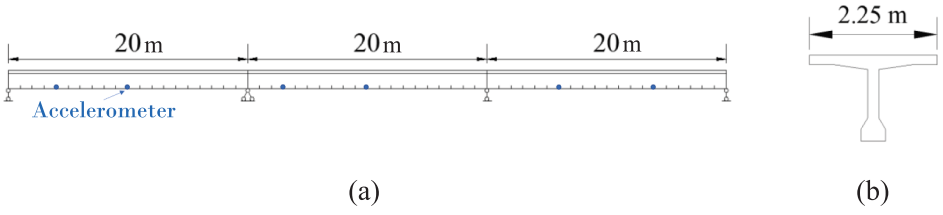


Fig. 1. (a) Layout of sensors on the bridge; (b) Layout of the cross section.

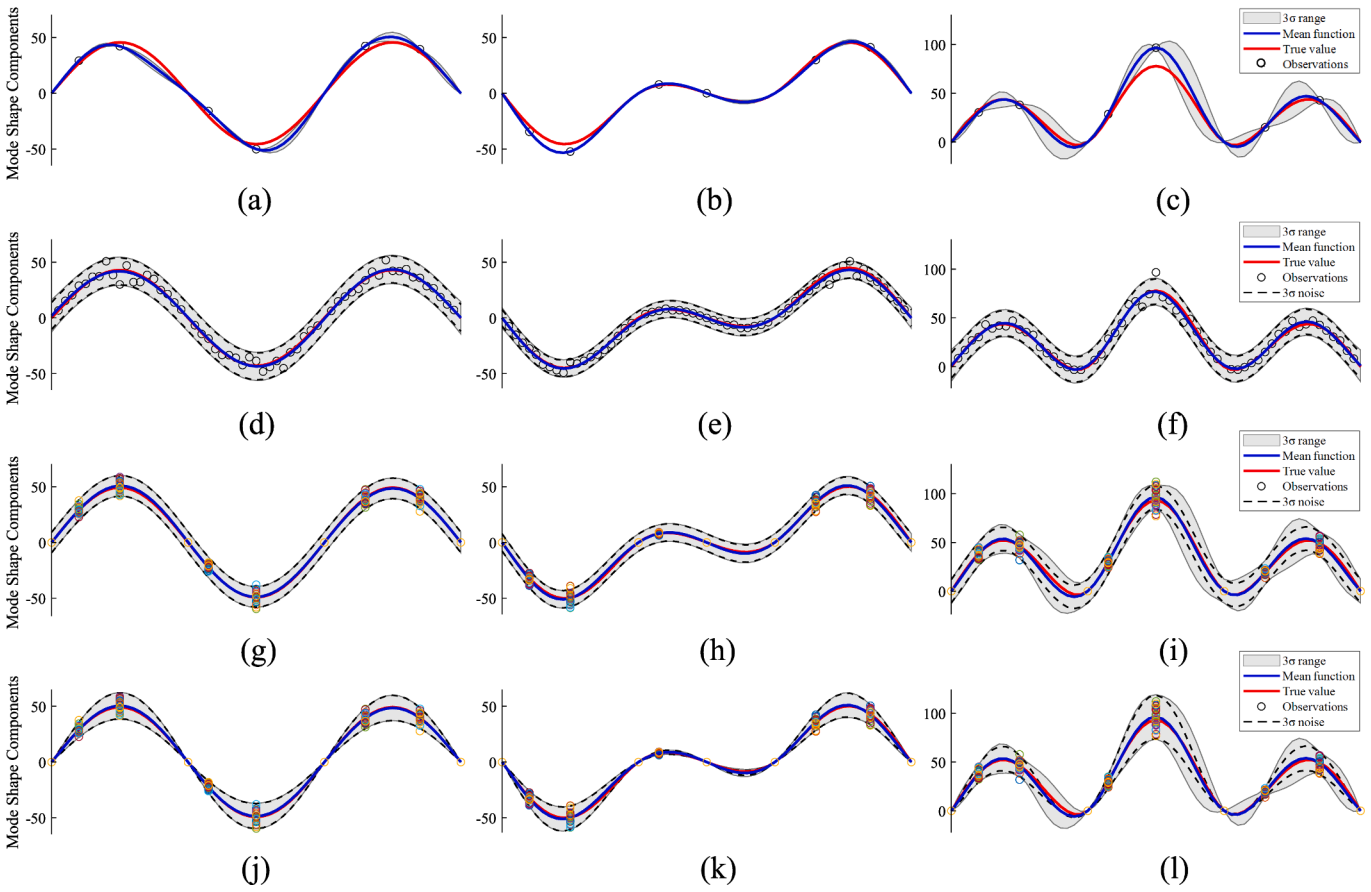


Fig. 2. (a-c) Results by GPR with 6 sensors; (d-f) Results by GPR with 61 sensors; (g-i) Results by RGPR with homoscedastic noise; (j-l) Results by RGPR with heteroscedastic noise.

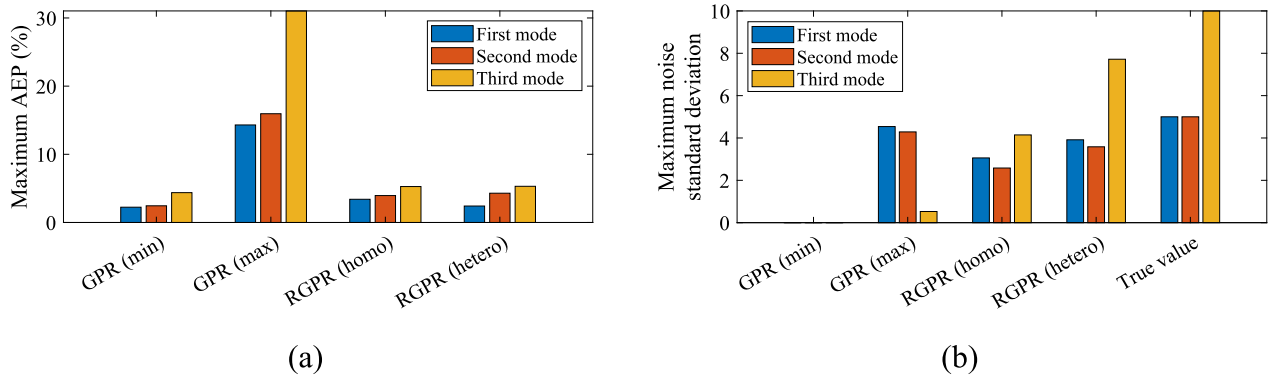


Fig. 3. Results comparison of (a) maximum AEP; (b) maximum noise standard deviation.

can be estimated and eliminated from the extracted modal information, and the modal parameters can finally be determined by some robust numerical techniques such as QR factorization and singular value decomposition. It is worth noting that the first few modes can generally be identified with satisfactory accuracy by OMA in real applications, so we suggest using the first several mode shapes in accordance with stably recognized frequencies in the sensor placement problem. Modes in high order are recommended only if they can be identified with high fidelity. Since temperature distribution may affect the identification results, 24 sets of time-domain acceleration signals, each lasting for one hour in nights, are selected. As was studied in [56], the 1st and 7th modes of this bridge are vertical vibration modes and show obvious difference in their curve shapes, so these two modes are tested in this study. Thus, 24 mode shape vectors from different time slots are obtained for each mode. The Guyan-reduced mass normalization method is employed to normalize the mode shape vectors [57,58]. Then the normalized mode shape components are scaled up to avoid pathologic calculation in GPR.

One of the 24 prediction models obtained by GPR and the prediction models obtained by homoscedastic and heteroscedastic RGPR for the two modes are shown in Fig. 5(a-c) and Fig. 6(a-c), respectively. The red lines in the detail drawings represent the 3σ range of true measurement noise, where the standard deviation is the unbiased estimation using 24 observations. Clearly, the noise standard deviation estimated by GPR for the 1st mode has a very small value which is hard to show in a visible way in Fig. 5(a), while the noise standard deviation estimated for the 7th mode has a quite large value. The reason is that the data quantity is too

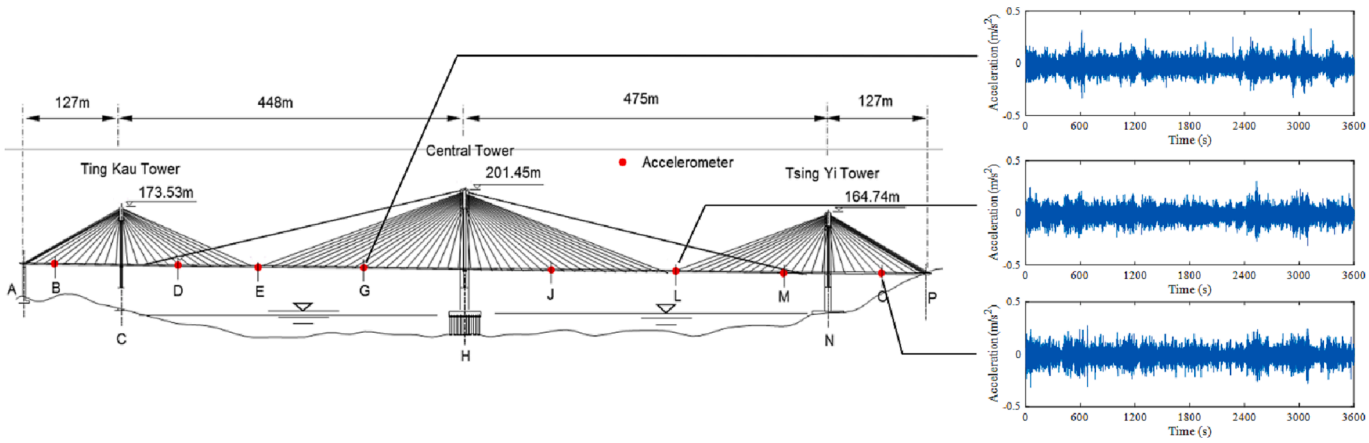


Fig. 4. Layout of Ting Kau bridge and accelerometers.

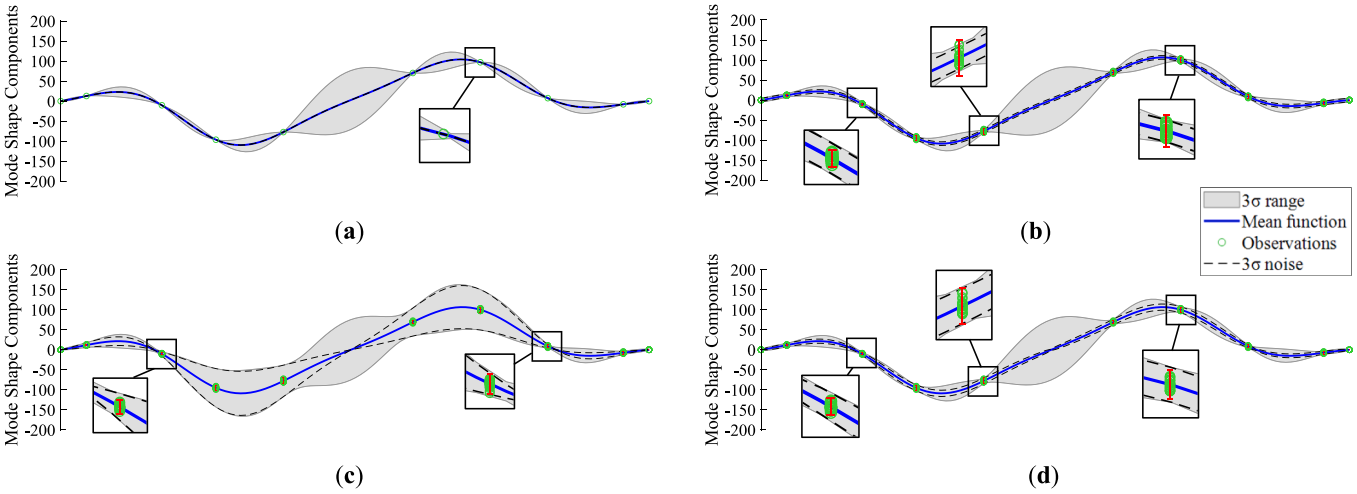


Fig. 5. Results of 1st mode by (a) GPR; (b) RGPR with homoscedastic noise; (c) RGPR with heteroscedastic noise (NRF $= f_p(\boldsymbol{x})$); (d) RGPR with heteroscedastic noise (NRF $= f_p^{1/4}(\boldsymbol{x})$).

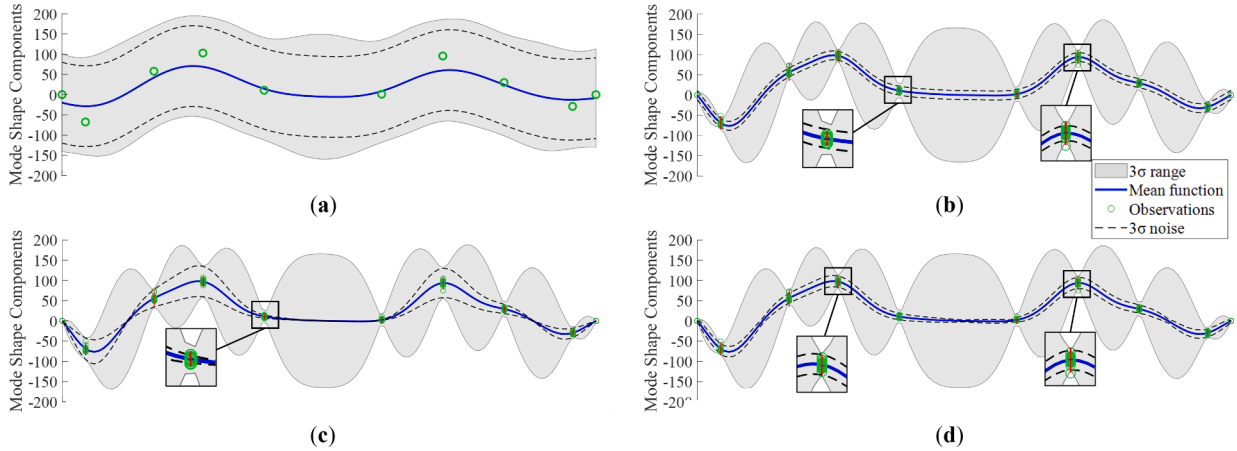


Fig. 6. Results of 7th mode by (a) GPR; (b) RGPR with homoscedastic noise; (c) RGPR with heteroscedastic noise (NRF $= f_p(\boldsymbol{x})$); (d) RGPR with heteroscedastic noise (NRF $= f_p^{1/4}(\boldsymbol{x})$).

small to support correct estimation of the noise in the modelling process. The noise variances estimated by homoscedastic RGPR do not match well with the true values at some of the sensed positions for both modes, and the noise standard deviation estimated by heteroscedastic RGPR has a large deviation from the true value. The reason might be that the standard deviation value of the noise is not in a linear relation with the mode shape value as assumed. In view of this, another assumption to NRF, $f_p^{1/4}(\boldsymbol{x})$, is tested. The results are shown in Fig. 5(d) and Fig. 6(d).

As is shown, the noise estimation results are much better now.

The maximum noise standard deviation values at the eight monitoring positions from different methods are compared with the true values calculated from real observations, as shown in Fig. 7. It can be observed that the maximum noise standard deviations estimated by the heteroscedastic RGPR model with NRF ${f_p}^{1/4}(\boldsymbol{x})$ are the closest to the real measurements. Even though GPR (min) shows a similar maximum noise standard deviation to the real measurements, the GPR method cannot provide stable estimations of the noise standard deviation (GPR (max) is much larger than the real measurements). In practical applications, we can modulate NRF according to the observations to improve modelling precision.

4. Two-stage strategy for globally optimal sensor arrangement

By using the heteroscedastic RGPR method, mode shape can be accurately reconstructed with a few deployed sensors, and information theory can be applied to instruct the sensor placement. As described in Section 1, it is preferred to directly use real observations to instruct the sensor placement so that the modelling error in using a finite element model can be avoided while measurement noise can be taken into account. However, deploying massive sensors to collect comprehensive structural responses is too expensive. To overcome this difficulty, a two-stage sensor placement strategy is proposed. The target of interest here is to precisely reconstruct mode shapes with as few sensors as possible. In the first stage, a few sensors are firstly deployed to collect data for training an accurate RGPR model as an alternative to the structure. The RGPR model can provide comprehensive structural response which is necessary for global optimization of sensor placement. The sensors in this stage are temporarily deployed, and the quantity and positions can be roughly determined. In the second stage, the sensors will be globally optimized in terms of quantity and positions with the target of accurate reconstruction of mode shapes. The temporarily deployed sensors in the first stage will be re-arranged to the optimum positions. In this way, the mode shape components will be precisely reconstructed with the minimum number of sensors.

In general, mode shapes of multiple modes are required for structural health monitoring, so the objective function in OSP should be designed for multiple modes. For example, the differential entropy values from multiple modes can be simply added to build the objective function, but the optimization balance among different modes couldn't be achieved. In the present study, this is achieved by scaling the mode shapes of different modes to the same level (e.g., the difference between the maximum value and minimum value of each mode shape is scaled up to the same value). When the RGPR prediction models are obtained, they can be scaled according to the mean function so that the difference between the maximum value and minimum value of each mode shape will be identical. In this light, the objective function can be formulated by adding together the differential entropies from different modes, and the optimization balance among different modes is expected to be attained.

4.1. The greedy algorithm for comprehensive information collection

In the first stage, the greedy algorithm is employed for optimization on account of its computation efficiency and operation simplicity [59]. An iteration process is built to simply add sensors one-by-one in sequence up to the needed amount. The deployment process is without retrieval, so it is practical for deployment on a real structure. Moreover, this algorithm will automatically terminate if the target of interest is achieved, thereby providing an upper limit for the best sensor quantity. This property can reduce the computational burden in the second-stage optimization in determining the best sensor quantity. One drawback of this algorithm is that the solution is probably near-optimal instead of optimal. However, this problem is negligible in this stage since the sensors are temporarily deployed and our target in this stage is to build an accurate RGPR model rather than attaining the final OSP solution. The flowchart of the algorithm combining with heteroscedastic RGPR is illustrated in Fig. 8.

A set of candidate points, which contains unsensed positions where sensors can be placed, will be defined before the optimization process. The initialized sensor scheme can be either newly deployed sensors for a bridge to be instrumented with an SHM system or the existing sensors already in a bridge where an SHM system has been installed but needs upgrade. They are suggested more than the number of target mode shapes so that these mode shapes can be distinguished from each other. We recommend that the initial sensor number is better to be larger than the target mode number by two to three. In each loop, the new RGPR prediction model is used to calculate the differential entropy at each candidate position. The position which has the largest entropy value will be deployed with a new sensor. Finally, the algorithm will be terminated if a predefined criterion is satisfied. In this study, the termination criterion is defined as

$$TC^{i+1} = \sum_{j=1}^{N_c}\left(H^{i}_{f_*(x_j)} - H^{i+1}_{f_*(x_j)}\right) \Big/ \sum_{j=1}^{N_c} H^{i}_{\varepsilon(x_j)} \tag{26}$$

in which i represents the number of iterations; j is the index of positions in the candidate set and N_c is the number of candidate positions; ε represents the measurement noise estimated by RGPR. This criterion calculates the ratio of uncertainty change of the prediction model during the iteration process to the uncertainty caused by the measurement noise. A small TC^{i+1} value implies that the uncertainty reduction in the $(i+1)$th iteration due to the newly added sensor is negligible compared to uncertainty caused by the measurement noise, and placing more sensors does not contribute to a reduction in uncertainty of the prediction model. Thus, when the change of the criterion between two consecutive loops is smaller than a given threshold, the iteration is

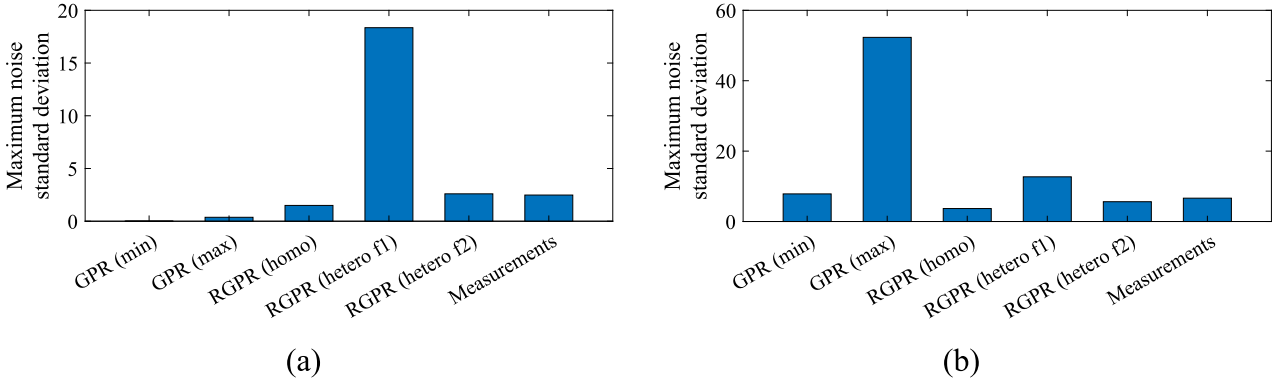


Fig. 7. Results comparison of maximum noise standard deviation: (a) 1st mode; (b) 7th mode.

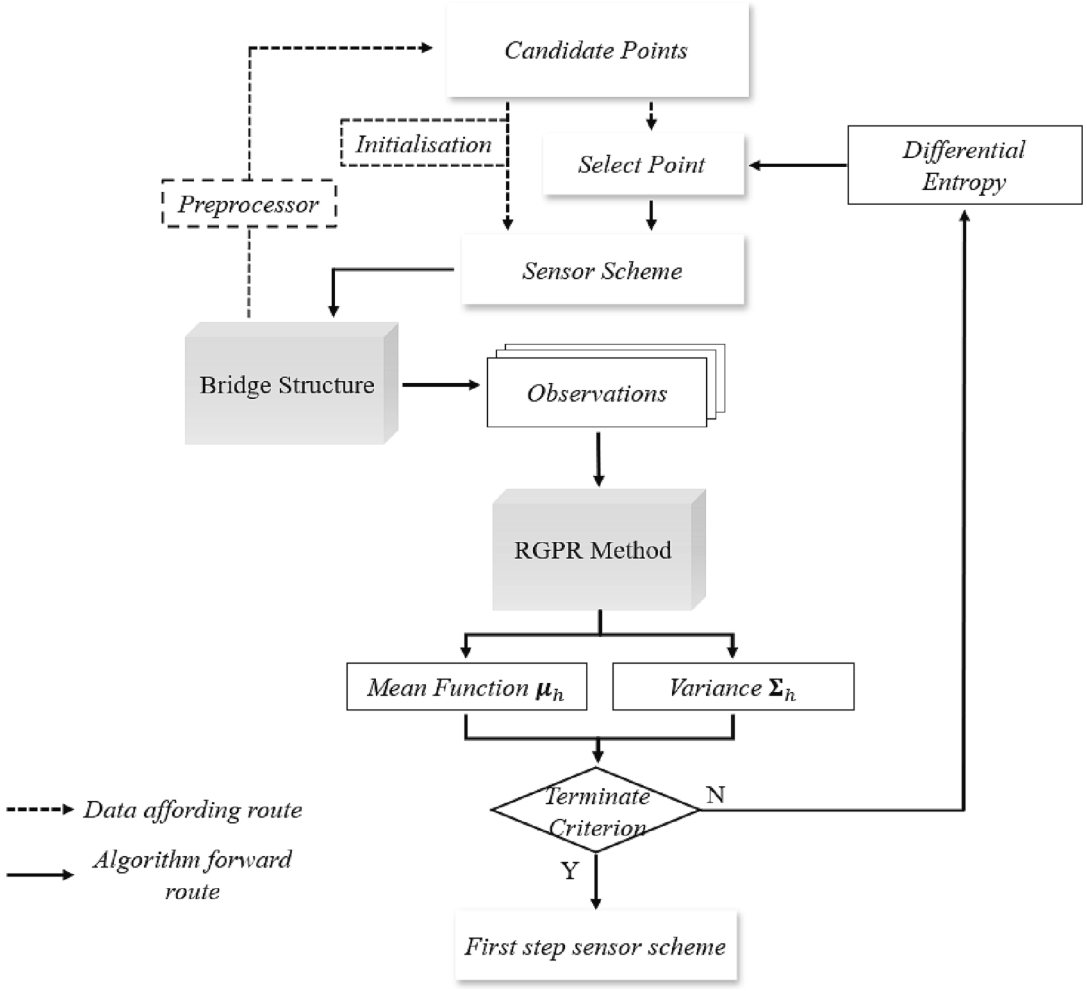


Fig. 8. Flowchart of the greedy algorithm with RGPR.

terminated. The detailed sensor configuration procedures are summarized as follows:

Algorithm 1. Greedy algorithm with heteroscedastic RGPR

Input: Candidate points $\mathscr{X}_N$; selection set S; observation set Y; initial sensor number N_0.

Initialization: Uniformly deploy initial sensors on a bridge and collect observations $Y = y(S)$; For a bridge already instrumented with an SHM system, directly collect observations from the sensor system.

In *i*th iteration:

1. Homoscedastic RGPR method to elicit prediction model and use the mean function as f_p;
2. Heteroscedastic RGPR method to derive the prediction model f_*;
3. Adjust NRF to make the estimated noise closer to the real observations;
4. If the termination criterion is satisfied, terminate the iteration; otherwise go to step 5;
5. Calculate the differential entropy at each point of the candidate set and choose the point that has the largest objective function value adding to the selection set $S = S \cup \boldsymbol{x}^{(i)}$; deploy sensor at this point and get observations $Y = Y \cup \boldsymbol{y}_{\boldsymbol{x}^{(i)}}$;
6. Update the candidate set by deleting the selected point $\mathscr{X}_N = \mathscr{X}_N \backslash \boldsymbol{x}^{(i)}$.

With the proposed algorithm, only a small number of sensors are needed to attain precise RGPR prediction models while information from a structural model is not required.

4.2. Cuckoo search algorithm for global optimization of sensor arrangement

With the precise RGPR prediction models formulated in the first stage, we can sample from the prediction models to obtain structural responses at unsensed locations according to the predicted mean function and the estimated noise. Such data, termed pseudo-observations, will be used to obtain the globally optimal sensor placement solution. A prominent heuristic algorithm, Cuckoo Search (CS), is employed as the optimization algorithm in this stage for sensor placement.

Like other heuristic algorithms, the CS algorithm is a random searching algorithm, which was proposed based on the parasitic brood behavior of some cuckoo species [60]. These kinds of cuckoos do not know how to build nest and brood for next generation, so they lay their eggs in the nests of other birds. Some of the host birds may bring them up. In an optimization view, the cuckoo's egg or the nest it located in is a feasible solution. The process of cuckoo searching nests for laying eggs represents a process to search for better feasible solution, and the survival of cuckoo means that the nest it located in is a good solution. Different from the traditional randomness in other heuristic algorithms, the CS algorithm adopts an advanced Lévy flight procedure, which has been demonstrated having good performance in global searching. As such, it affords high calculation efficiency and can prevent the algorithm from falling into local optimal solutions. Besides, it has a simple algorithm structure and is easy to implement compared with other heuristic algorithms. The details of the CS algorithm with Lévy flight can refer to [61].

In the CS algorithm, the objective function is defined different from what one uses in the greedy algorithm, which is the differential entropy of an individual point as shown in Equation (10). A Global Differential Entropy (GDE) is proposed, which is more suitable for global optimization:

$$GDE = -\int\int p(f_*(\boldsymbol{x}_*))lnp(f_*(\boldsymbol{x}_*))df_*d\boldsymbol{x}_* = \int \frac{1}{2}(1+ln2\pi+ln|\boldsymbol{\Sigma}(\boldsymbol{x}_*)|)d\boldsymbol{x}_* \quad (26)$$

The GDE reflects the uncertainty of an obtained prediction model instead of an individual point. For a specific sensor scheme, a prediction model along with its GDE value will be obtained. The sensor scheme with the minimum GDE value denotes that the obtained prediction model has the minimum uncertainty. As such, our target is to choose a sensor scheme which minimizes the GDE. However, Equation (26) is analytically intractable. As an alternative, numerical approximation by calculating discrete differential entropy values at some representative points is adopted:

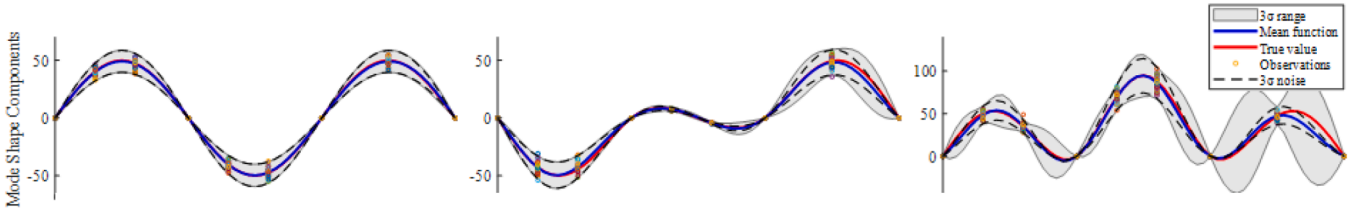
$$GDE \approx \frac{1}{2}\sum_{i=1}^{N_r}\Delta\boldsymbol{x}_{*i}(1+ln2\pi+ln|\boldsymbol{\Sigma}(\boldsymbol{x}_{*i})|) \quad (27)$$

in which i and N_r denote the index and total number of representative points, respectively; and $\Delta\boldsymbol{x}_{*i} = \boldsymbol{x}_{*i} - \boldsymbol{x}_{*(i-1)}$. The representative points are user-defined, which can be either the integer values of coordinate or the nodes in the finite element model.

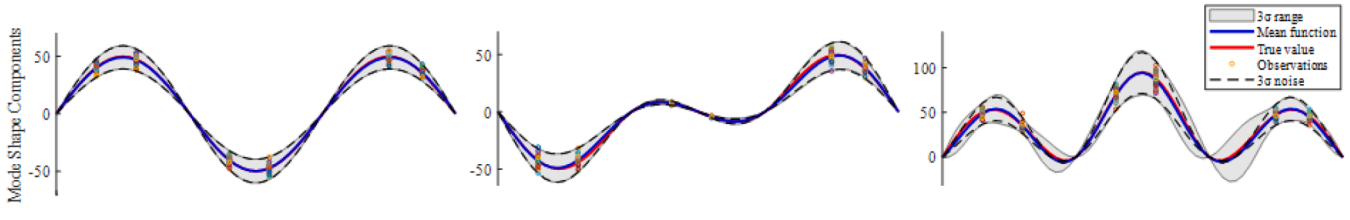
In the CS algorithm, the sensor number is one of the inputs, so it should be determined before the calculation. It can be approximately estimated by referring to the sensor quantity obtained in the first stage. Because the greedy algorithm searches for near-optimal solution, the sensors considered in the first stage would be redundant if they are optimally re-arranged. Hence, we can test different sensor numbers that are smaller than the sensor quantity elicited in the first stage. By comparing the results with different numbers of sensors, the best sensor quantity and configuration can be achieved. Moreover, several parameters in the CS algorithm should be assigned proper values before pursuing the algorithm, such as the discovery probability, nest number and maximum iterations. The values of these parameters are chosen in this study according to the recommendations given in [61].

5. Case study

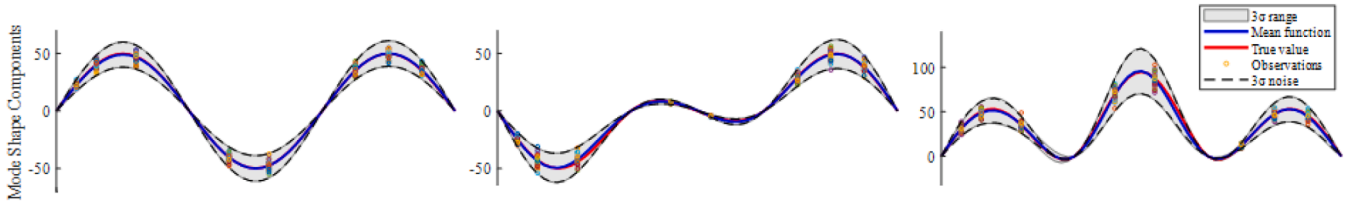
To evaluate the proposed two-stage sensor placement strategy, the numerical three-span continuous beam bridge model which has been introduced in Section 3.1 is explored. The mode shapes of the first three modes are concerned.



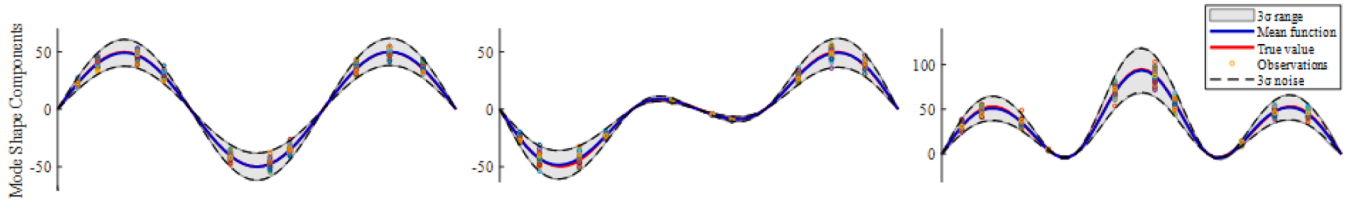
(a) 1st iteration with 5 sensors



(b) 2nd iteration with 6 sensors



(c) 4th iteration with 8 sensors



(d) 6th iteration with 10 sensors

Fig. 9. Prediction models of three mode shapes during iteration process.

5.1. First stage optimization

As initialization, five accelerometers are uniformly deployed on the beam bridge model, which can either be those from an existing monitoring system or the initialized scheme of a new monitoring system, and the mode shape values at these five positions are made available. Heteroscedastic measurement noise, whose standard deviation is assumed as linear to the mode shape value, is artificially added. The noise standard deviation is 10 % of the mode shape magnitude. The criterion threshold for iteration termination is set as 5 %. The obtained prediction models during the iteration process are illustrated in Fig. 9.

As can be seen from the panels in this figure, when five sensors are deployed on the bridge, the first mode shape can be accurately modelled, but the second and third mode shapes have large prediction uncertainties at some positions. With the adding of sensors, the prediction uncertainties of the second and third mode shapes gradually decrease. Finally, the algorithm is terminated with 10 sensors, and the models for the three mode shapes all have low prediction uncertainties. Moreover, with the increase of sensors, the mean functions of the three prediction models are increasingly close to the true values. The change of the termination criterion value during the iteration process is depicted in Fig. 10, which becomes steady when the number of iterations exceeds five. The GDE values of the three mode shapes versus sensor quantity are displayed in Fig. 11, which illustrates the uncertainty of the prediction models. The maximum AEP with respect to sensor quantity is also depicted to show the accuracy of the predicted mean functions.

As illustrated in Fig. 11, the maximum AEP values of the second and third mode shapes in general decrease with increasing number of sensors, demonstrating that the precision of the prediction models of the second and third mode shapes increases gradually with the adding of sensors. On the other hand, the maximum AEP values of the first mode shape always have the lowest values among the three mode shapes, so fewer sensors are required for accurate modelling of the first mode shape than the second and third mode shapes. Besides, it can be found that the GDE values show similar trends with the AEP values. This tallies with the observation made in Fig. 10 that the predicted mean functions will be closer to the system true values when the prediction uncertainty becomes smaller. Therefore, GDE can be regarded as a reasonable alternative of AEP to evaluate the RGPR model when applying the proposed method in real applications. The sensor placement results are shown in Fig. 12.

By using the predicted mean function and estimated noise, we can obtain the values of the three mode shapes at any point on the structure. These pseudo-observations are regarded as alternatives to the real structural responses, which can be employed for global optimization of sensor arrangement in the next stage.

5.2. Second stage optimization

In the second stage, the sensor quantity as one of the inputs in the CS algorithm is first determined. As indicated by the results from the first stage, the best sensor quantity will not exceed 10. In view of this, we test three sensor quantities, 8, 9 and 10, and choose the best sensor quantity by comparing their optimization results. The evolution processes of the CS algorithm under three sensor quantities are plotted in Fig. 13. As can be seen, the algorithm finally converges at the end of the iteration process under three sensor quantities, which means that the best sensor configuration schemes are obtained. The GDE of the prediction models with 8 sensors finally converges to a value larger than those with 9 sensors and 10 sensors. For clear comparison, the GDE values from the CS algorithm are compared with the results obtained by the greedy algorithm with 9 and 10 sensors, as depicted in Fig. 14. It can be observed that the sensor configuration schemes for 9 and 10 sensors optimized by the CS algorithm can help train the prediction models with GDE values similar to those obtained by the greedy algorithm with 10 sensors. Decreasing to 8 sensors will cause a big increase of prediction uncertainty of the RGPR models, especially the models for the first and third modes. Hence, it can be concluded that the RGPR models do not possess desired accuracy with data collected from 8 sensors. Besides, the GDE of 9 sensors using the greedy algorithm is larger than that of 9 sensors using the CS algorithm, which demonstrates that the second-stage optimization can get better sensor configuration result than that from the first stage. Therefore, the best sensor quantity in this case is 9, and the corresponding configuration scheme is the optimal result.

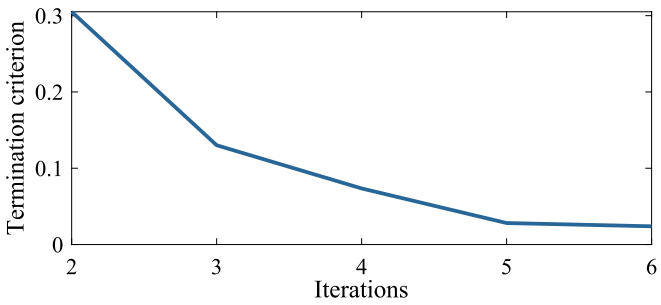


Fig. 10. Termination criterion.

To validate the accuracy of the obtained results, the AEPs of the prediction models obtained by the greedy algorithm and the CS algorithm are shown in Figs. 15 and 16, respectively. Through comparison, it is obvious that the results of 9 and 10 sensors optimized by the CS algorithm show similar accuracy to those of 10 sensors using the greedy algorithm while much better performance than those of 8 sensors. This observation in accordance with the GDE results demonstrates the effectiveness of using GDE in optimization. In addition, the AEP results of 9 sensors optimized by the CS algorithm are smaller than those from 9 sensors using the greedy method, which validates that the second-stage optimization is effective.

Finally, the prediction models with 9 sensors are portrayed in Fig. 17, and the sensor configuration under 9 sensors is shown in Fig. 18. By comparing the modelling results in Fig. 17 with those in Fig. 9(a), it is apparent that the optimized sensor scheme can provide much higher modelling accuracy than the original sensor scheme, especially the second and third modes. It is therefore demonstrated that the proposed two-stage strategy is effective in application to the existing monitoring system for upgrading.

6. Conclusions

This paper proposes a data-driven two-stage Optimal Sensor Placement (OSP) strategy aiming at the reconstruction of mode shapes for bridge structures. A novel Recurrent Gaussian Process Regression (RGPR) approach is presented for this purpose. By utilizing data from the time-domain signals to improve the estimation accuracy of measurement noise, the proposed RGPR approach can model the spatial mode shapes accurately with using only a small number of sensors. In line with this approach, the first-stage OSP deploys a few temporary sensors to collect data for training precise RGPR prediction models. These models can be regarded as alternatives to the real structure for generating structural responses at unsensed positions. Then the second-stage OSP employs the Cuckoo Search (CS) algorithm to attain the globally optimal sensor configuration, in which the required data is generated by the RGPR models. The temporarily deployed sensors are re-arranged to the optimum positions in the second stage.

The proposed RGPR approach is validated on a simulated three-span continuous bridge model and the cable-stayed Ting Kau bridge. Results show that the RGPR approach can obtain more accurate prediction models than the GPR method. The efficiency of the two-stage OSP strategy is verified on the simulation model. It is demonstrated that the best sensor quantity and sensor placement configuration can both be obtained by the proposed strategy. In summary, the proposed two-stage

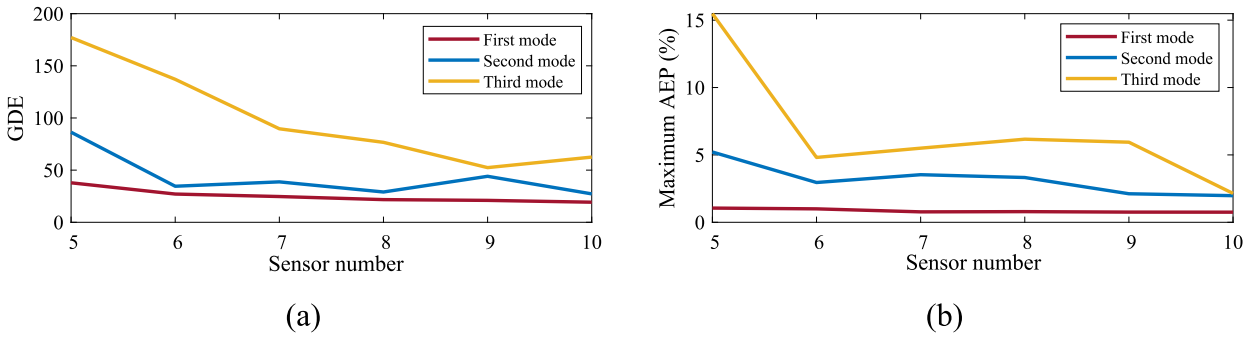


Fig. 11. (a) GDE and (b) Maximum AEP.

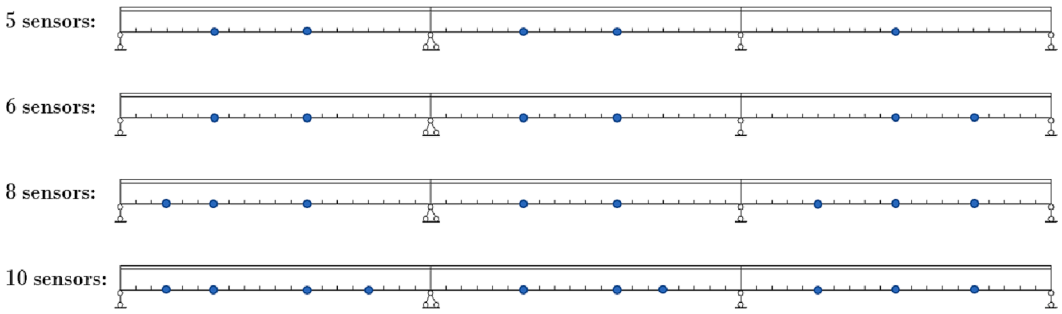


Fig. 12. Sensor configurations during iteration process.

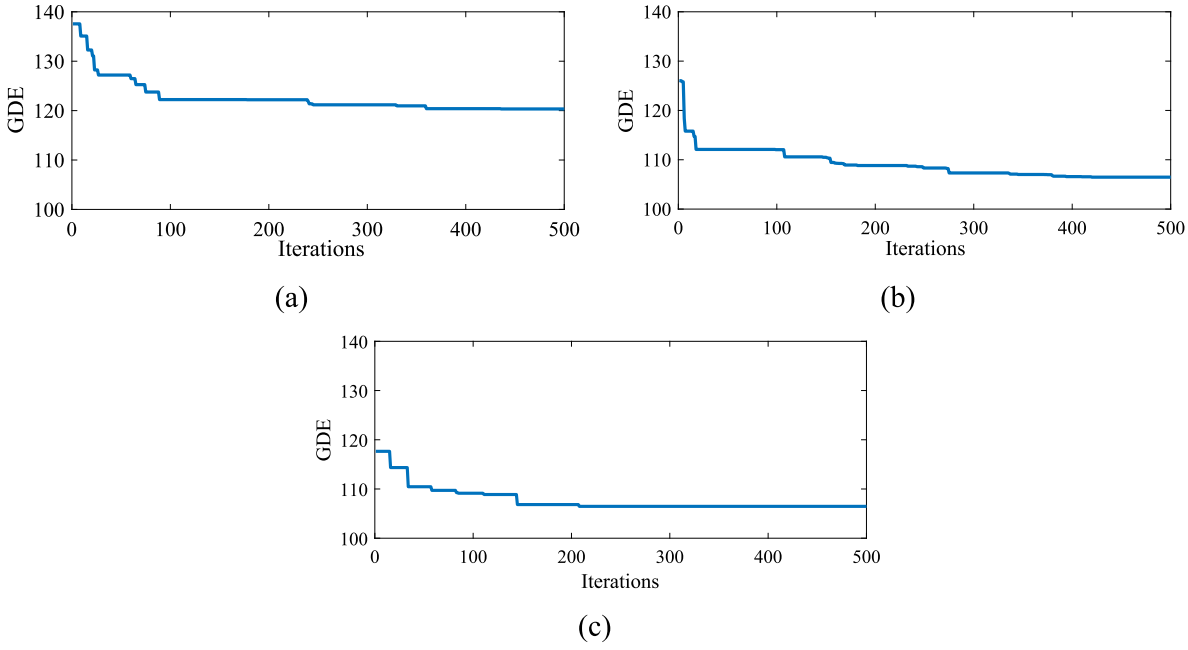


Fig. 13. Evolution process under (a) 8 sensors; (b) 9 sensors; (c) 10 sensors.

OSP strategy with the RGPR method has the following advantages: First, comprehensive structural information can be obtained by deploying a few sensors on a bridge, so the test cost is low. Second, the sensor configuration process is fully data-driven, so the modelling error is eliminated. Third, the measurement noise can be modelled with either a homoscedastic or heteroscedastic model, and it can be taken into account in the OSP process. Fourth, both sensor positions and sensor quantity can be optimized.

The proposed OSP strategy does not rely on a finite element model, so the result is more feasible for real structures than the traditional OSP methods. In the future, the concept of data-driven OSP strategy will be combined with other sensor configuration approaches to achieve more practical sensor arrangement results. For example, the finite element model can be updated by using the reconstructed mode shapes in the first stage, and then various objective functions can be formulated in the second stage with the use of information from the refined finite element model.

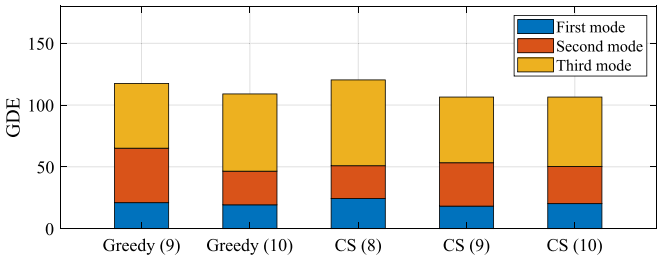


Fig. 14. Comparison of GDE.

CRediT authorship contribution statement

Bei-Yang Zhang: Methodology, Data curation, Software, Formal analysis, Writing – original draft, Visualization. **Yi-Qing Ni:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

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

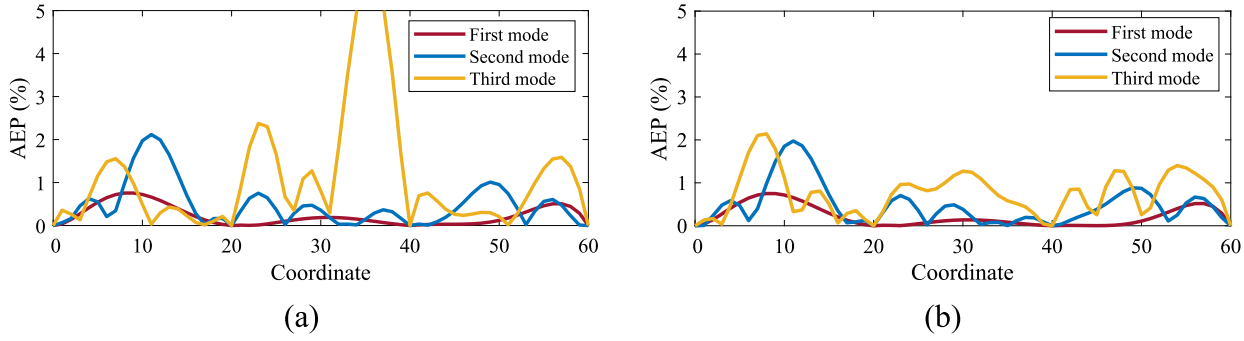


Fig. 15. AEP of prediction models by the greedy algorithm under (a) 9 sensors; (b) 10 sensors.

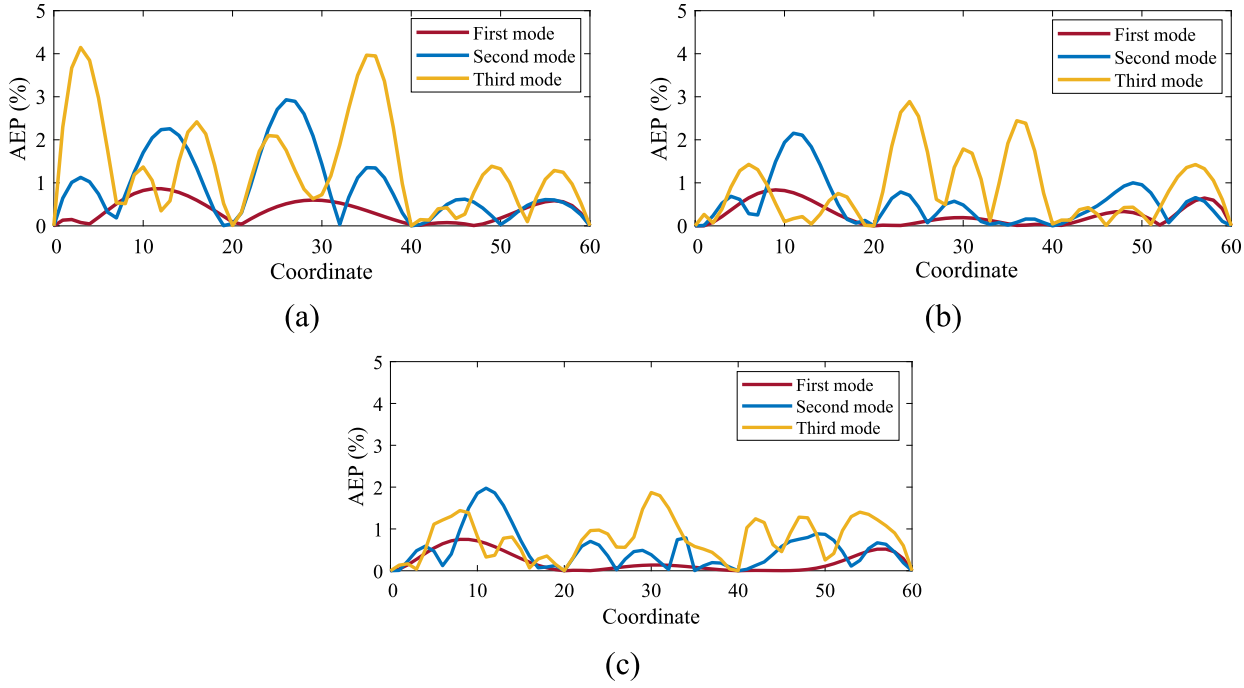


Fig. 16. AEP of prediction models by the CS algorithm under (a) 8 sensors; (b) 9 sensors; (c) 10 sensors.

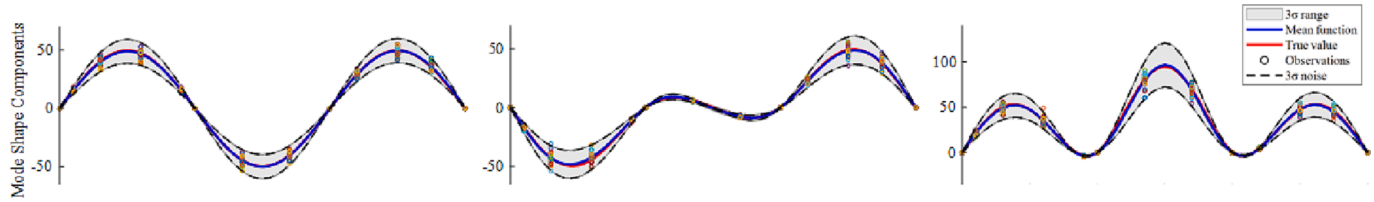


Fig. 17. Prediction models with 9 sensors.

Fig. 18. Optimal sensor configuration under 9 sensors.

Data availability

Data will be made available on request.

Acknowledgements

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (SAR), China (Grant No. PolyU 152014/18E) and a grant from the Guangdong Basic and Applied Basic Research Foundation of Department of Science and Technology of Guangdong Province (Grant No. 2021B1515130006). The authors also appreciate the funding support by the Innovation and Technology Commission of Hong Kong SAR Government to the Hong Kong Branch of Chinese National Rail Transit Electrification and Automation Engineering Technology Research Center (Grant No. K-BBY1).

References

[1] Yu S, Ou J. Structural health monitoring and model updating of Aizhai suspension bridge. J Aerosp Eng 2017;30:B4016009.
[2] Yang C. Sensor placement for structural health monitoring using hybrid optimization algorithm based on sensor distribution index and FE grids. Struct Control Heal Monit 2018;25:e2160.
[3] Gao Q, Cui K, Li J, Guo B, Liu Y. Optimal layout of sensors in large-span cable-stayed bridges subjected to moving vehicular loads. Int J Distrib Sens Networks 2020;16:1–13.
[4] Liu C, Jiang Z, Gong Y, Xiao Y. A two-stage optimal sensor placement method for multi-type structural response reconstruction. Meas Sci Technol 2021;32:035114.
[5] Kammer DC. Sensor placement for on-orbit modal identification and correlation of large space structures. J Guid Control Dyn 1991;14:251–9.
[6] Heo G, Wang ML, Satpathi D. Optimal transducer placement for health monitoring of long span bridge. Soil Dyn Earthq Eng 1997;16:495–502.
[7] Kwon SJ, Shin S, Lee HS, Park YH. Design of accelerometer layout for structural monitoring and damage detection. KSCE J Civ Eng 2003;7:717–24.
[8] Li DS, Li HN, Fritzen CP. A note on fast computation of effective independence through QR downdating for sensor placement. Mech Syst Signal Process 2009;23:1160–8.
[9] Kim T, Youn BD, Oh H. Development of a stochastic effective independence (SEFI) method for optimal sensor placement under uncertainty. Mech Syst Signal Process 2018;111:615–27.
[10] Meo M, Zumpano G. On the optimal sensor placement techniques for a bridge structure. Eng Struct 2005;27:1488–97.
[11] Ewins DJ. Modal testing: theory and practice. Letchworth: Research Studies Press Ltd; 1984.
[12] Yi TH, Li HN, Gu M. Optimal sensor placement for health monitoring of high-rise structure based on Genetic Algorithm. Math Probl Eng 2011:395101.
[13] Zhang BY, Zhang XD, Zhang ZH. Nested-stacking genetic algorithm for the optimal placement of sensors in bridge. Comput J 2018;61:1269–83.
[14] Yin H, Dong K, Pan A, Peng Z, Jiang Z, Li S. Optimal sensor placement based on relaxation sequential algorithm. Neurocomputing 2019;344:28–36.
[15] Papadimitriou C, Beck JL, Au SK. Entropy-based optimal sensor location for structural model updating. J Vib Control 2000;6:781–800.
[16] Papadimitriou C. Optimal sensor placement methodology for parametric identification of structural systems. J Sound Vib 2004;278:923–47.
[17] Yin T, Yuen KV, Lam HF, Zhu H. Entropy-based optimal sensor placement for model identification of periodic structures endowed with bolted joints. Comput Civ Infrastruct Eng 2017;32:1007–24.
[18] Yin T, Zhang FL. Sensor placement for model identification of multi-story buildings under unknown earthquake ground motion. Eng Struct 2022;251:113548.
[19] Mallardo V, Aliabadi MH. Optimal sensor placement for structural, damage and impact identification: a review. Struct Durab Heal Monit 2013;9:287–323.
[20] Cobb RG, Liebst BS. Sensor placement and structural damage identification from minimal sensor information. AIAA J 1997;35:369–74.
[21] Shi Q, Wang X, Chen W, Hu K. Optimal sensor placement method considering the importance of structural performance degradation for the allowable loadings for damage identification. Appl Math Model 2020;86:384–403.
[22] Worden K, Burrows AP. Optimal sensor placement for fault detection. Eng Struct 2001;23:885–901.
[23] Shi ZY, Law SS, Zhang LM. Optimum sensor placement for structural damage detection. J Eng Mech 2000;126:1173–9.
[24] Guo YL, Ni YQ, Chen SK. Optimal sensor placement for damage detection of bridges subject to ship collision. Struct Control Heal Monit 2017;24:e1963.
[25] Guo HY, Zhang L, Zhang LL, Zhou JX. Optimal placement of sensors for structural health monitoring using improved genetic algorithms. Smart Mater Struct 2004;13:528–34.
[26] Lenticchia E, Ceravolo R, Chiorino C. Damage scenario-driven strategies for the seismic monitoring of XX century spatial structures with application to Pier Luigi Nervi's Turin Exhibition Centre. Eng Struct 2017;137:256–67.
[27] Ye SQ, Ni YQ. Information entropy based algorithm of sensor placement optimization for structural damage detection. Smart Struct Syst 2012;10:443–58.
[28] Ye SQ, Ni YQ. Sensor placement strategy for structural damage detection from modal strain energy change. In Proceedings of the 14th Asia Pacific Vibration Conference, 5-8 December 2011, Hong Kong SAR, P.R. China;2011:141–6.
[29] Zare Hosseinzadeh A, Ghodrati Amiri G, Seyed Razzaghi SA, Koo KY, Sung SH. Structural damage detection using sparse sensors installation by optimization procedure based on the modal flexibility matrix. J Sound Vib 2016;381:65–82.
[30] Zhou K, Wu ZY. Strain gauge placement optimization for structural performance assessment. Eng Struct 2017;141:184–97.
[31] Zhang XH, Zhu S, Xu YL, Homg XJ. Integrated optimal placement of displacement transducers and strain gauges for better estimation of structural response. Int J Struct Stab Dyn 2011;11:581–602.
[32] Lin JF, Xu YL, Law SS. Structural damage detection-oriented multi-type sensor placement with multi-objective optimization. J Sound Vib 2018;422:568–89.
[33] Zhu S, Zhang XH, Xu YL, Zhan S. Multi-type sensor placement for multi-scale response reconstruction. Adv Struct Eng 2013;16:1779–97.
[34] Xu YL, Zhang XH, Zhu SY, Zhan S. Multi-type sensor placement and response reconstruction for structural health monitoring of long-span suspension bridges. Sci Bull 2016;61:313–29.
[35] Li ZN, Tang J, Li QS. Optimal sensor locations for structural vibration measurements. Appl Acoust 2004;65:807–18.
[36] Guestrin C, Krause A, Singh AP. Near-optimal sensor placements in Gaussian processes. In Proceedings of the 22nd international conference on Machine learning, 7-11 August 2005, Bonn, Germany;2011:265–72.
[37] Krause A, Guestrin C, Gupta A, Kleinberg J. Near-optimal sensor placements: maximizing information while minimizing communication cost. In Proceedings of the 5th International Conference on Information Processing in Sensor Networks, 19-21 April, 2006, Nashville U.S. America;2006:2–10.
[38] Krause A, Singh A, Guestrin C. Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies. J Mach Learn Res 2008;9:235–84.
[39] He C, Xing J, Li J, Yang Q, Wang R, Zhang X. A combined optimal sensor placement strategy for the structural health monitoring of bridge structures. Int J Distrib Sens Networks 2013;2013:820694.
[40] Chattopadhyay A, Seeley CE. A simulated annealing technique for multiobjective optimization of intelligent structures. Smart Mater Struct 1994;3:98–106.
[41] Jia J, Feng S, Liu W. A triaxial accelerometer monkey algorithm for optimal sensor placement in structural health monitoring. Meas Sci Technol 2015;26:065104.
[42] Zhao J, Wu X, Sun Q, Zhang L. Optimal sensor placement for a truss structure using particle swarm optimisation algorithm. Int J Acoust Vib 2017;22:439–47.
[43] Yi TH, Li HN. Methodology developments in sensor placement for health monitoring of civil infrastructures. Int J Distrib Sens Networks 2012:612726.
[44] Sun H, Büyüköztürk O. Optimal sensor placement in structural health monitoring using discrete optimization. Smart Mater Struct 2015;24:125034.
[45] Zhou GD, Xie MX, Yi TH, Li HN. Optimal wireless sensor network configuration for structural monitoring using automatic-learning firefly algorithm. Adv Struct Eng 2019;22:907–18.
[46] Murugan Jaya M, Ceravolo R, Zanotti Fragonara L, Matta E. An optimal sensor placement strategy for reliable expansion of mode shapes under measurement noise and modelling error. J Sound Vib 2020;487:115511.
[47] Lin TY, Tao J, Huang HH. A multiobjective perspective to optimal sensor placement by using a decomposition-based evolutionary algorithm in structural health monitoring. Appl Sci 2020;10:7710.
[48] Pachón P, Infantes M, Cámara M, Compán V, García-Macías E, Friswell MI, et al. Evaluation of optimal sensor placement algorithms for the structural health monitoring of architectural heritage: application to the monastery of San Jerónimo de Buenavista (Seville, Spain). Eng Struct 2020;202:109843.
[49] Manohar K, Brunton BW, Kutz JN, Brunton SL. Data-driven sparse sensor placement for reconstruction: demonstrating the benefits of exploiting known patterns. IEEE Control Syst 2018;38:63–86.
[50] Xu Y, Choi J, Oh S. Mobile sensor network navigation using Gaussian processes with truncated observations. IEEE Trans Robot 2011;27:1118–31.
[51] Zhou HF, Ni YQ, Ko JM. Structural health monitoring of the Jiangyin Bridge: system upgrade and data analysis. Smart Struct Syst 2013;11:637–62.
[52] Rasmussen CE, Williams CKI. Gaussian processes for machine learning. London: the MIT Press; 2006.
[53] Cover TM, Thomas JA. Elements of information theory. 2nd ed. Hoboken: John Wiley & Sons; 2005.
[54] Tondreau G, Reynders E, Deraemaeker A. Towards a more realistic modelling of the uncertainty on identified mode shapes due to measurement noise. J Phys Conf Ser 2011;305:012002.
[55] Bergermann R, Schlaich M. Ting Kau bridge, Hong Kong. Eng Struct 1996;6:152–4.

[56] Ni YQ, Wang YW, Xia YX. Investigation of mode identifiability of a cable-stayed bridge: comparison from ambient vibration responses and from typhoon-induced dynamic responses. Smart Struct Syst 2015;15:447–68.
[57] Doebling SW, Farrar CR. Computation of structural flexibility for bridge health monitoring using ambient modal data. In Proceedings of 11th ASCE Engineering Mechanics Conference, 19-22 May, 1996, Fort Lauderdale, U.S. America;1996: 1114–7.
[58] Guyan RJ. Reduction of stiffness and mass matrices. AIAA J 1965;3:380.
[59] Manohar K, Kutz JN, Brunton SL. Optimal sensor and actuator selection using balanced model reduction. IEEE Trans Automat Contr 2022;67:2108–15.
[60] Yang XS. Nature-inspired metaheuristic algorithms. 2nd ed. Frome: Luniver Press; 2010.
[61] Yang XS, Deb S. Engineering optimisation by Cuckoo search. Int J Math Model Numer Optim 2010;1:330.