This is a repository copy of *Bayesian data driven model for uncertain modal properties identified from operational modal analysis*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/153802/

Version: Published Version

Article:

Bayesian data driven model for uncertain modal properties identified from operational modal analysis

Yi-Chen Zhu [a,*], Siu-Kui Au [b]

[a] *Dynamics Research Group, Department of Mechanical Engineering, The University of Sheffield, UK*
[b] *School of Civil & Environmental Engineering, Nanyang Technological University, Singapore*

A R T I C L E I N F O

A B S T R A C T

In structural health monitoring (SHM), 'data driven models' are often applied to investigate the relationship between the dynamic properties of a structure and environmental/operational conditions. Dynamic properties and environmental/operational conditions may not be directly measured but are rather inferred based on measured structural response data. Conventional data driven models assume training data as precise values without uncertainty, but this may not be justified when they are identified by operational modal analysis (OMA) where identification uncertainty can be significant. The associated confidence or precision may also vary depending on their identification uncertainties. This paper develops a Bayesian data driven model for modal properties identified from OMA. Identification uncertainty is incorporated fundamentally through the posterior distribution of modal properties of interest given the ambient vibration measurements. A Gaussian Process model is used for describing the potential unknown relationship between the modal properties and environmental/operational condition, which is subjected to OMA identification uncertainty. An efficient framework is developed to facilitate computation. The proposed method is validated by synthetic and laboratory data. Typhoon data from two tall buildings illustrates the field application of the proposed method.

1. Introduction

Structural health monitoring (SHM) aims at assessing the physical conditions of structures based on measured response data with applications in, e.g., damage detection and maintenance [1–5]. With an increasing number of SHM systems built into modern structures such as tall buildings and long-span bridges [1,6], it has the premise for evaluating structural serviceability and reliability throughout their whole life-cycle. Various quantities can be measured for SHM including strain, pressure and vibration. Among others, vibration-based SHM technique has become a popular non-destructive method where the modal parameters (e.g., natural frequencies, damping ratios and mode shapes) identified from vibration response data are investigated [7–10]. One premise is that by tracking the changes of these modal parameters against the environmental or operational conditions, it may be possible to monitor the health condition of the subject structure.

Operational modal analysis (OMA), also known as ambient modal identification, aims at identifying modal parameters based on vibration response of a tested structure under natural excitations such as wind, microtremor and cultural activities. It can be conducted when the structure is under operational condition without artificial loading, where the unknown

---

excitations are assumed to be statistically random. For its high economy and convenience in applications, OMA has gained popularity over the past few decades [11–14]. Since the loading information is unknown and cannot be directly controlled, the associated uncertainty of identified modal parameters is a major concern in OMA. Methods have been developed for quantifying identification uncertainties, e.g. [15–17].

One way of understanding dynamic behaviour of structures under different conditions is to apply 'data driven regression models' [18–20], which aim at expressing the identified modal parameter values of structures from OMA as a function of environmental or operational variables. In this regard, the modal properties and the environmental/ operational variables are taken as training data for constructing the data driven models. Models used in the literature include Polynomial Chaos Expansions [21]; Functionally Pooled model [22] and Kernel Principal Component Analysis [23]. Among others, Gaussian Process (GP) has been found to offer an effective means for constructing data driven models [24]. Applications include using GP time-series model with Principal Component Regression for structural response prediction [25]; treed GP model for switching different environmental regimes [20]; GP model with the generalized likelihood ratio test for novelty detection [26] and GP model under Engle-Granger framework for nonlinear cointegration [27]. A typical assumption of conventional data driven models is that the training data are known precisely. However, this is usually not the case in SHM. For vibration-based SHM, both the modal parameters and environmental/operational conditions may not be directly measured but are rather identified from measured response data, which inevitably carry imprecision. Such uncertainty can be affected by test configurations such as sensor noise and measurement duration, which can also vary among the identified training data points. GP models considering uncertain input training data have been developed [28] where uncertainty is effectively incorporated when calculating the expected covariance function. It was shown that the expected covariance function can be obtained analytically when using the squared exponential function and assuming Gaussian input noise. Monte Carlo expectation maximization method has also been applied to compute the marginal evidence under uncertain input data in training Gaussian process models [29]. The general problem of training a data driven model with uncertain input has been studied in [30], which is known as 'errors-in-variables regression'.

Conventionally, the uncertainty in output data is considered through the additive modelling error term in the GP model. This is different from the imprecision arising from the identification uncertainty of training data in the SHM context of this work, where the uncertainty need not be additive and the distribution should be consistent with its origin (e.g., how modal properties are identified) rather than being subjected to heuristic choice.

Motivated by the above considerations, this paper proposes a Bayesian framework capable of inferring data driven models with training data identified from OMA where their associated identification uncertainty is incorporated through the posterior distribution given the ambient vibration measurements. Efficient computation requires one to express the posterior distribution of hyper parameters (related to data driven model) given the ambient vibration measurements in an explicit form but that is highly non-trivial. In Section 3 we accomplish this for a general context (i.e., not limited to OMA) following a Bayesian approach, where the posterior distribution of hyper parameters is expressed in terms of that of the modal properties (training data) and the marginal distribution of output training data given the input training data as well as the hyper parameters. Theoretical issues are investigated in detail. The resulting formula is intuitive and conducive to analysis and computation. The general framework is specialised to OMA and efficient computation strategy is proposed in Section 4 where a GP model is used for inferring the hyperparameters. Computational efficiency is further enhanced in Section 5 by considering the characteristics of OMA identification uncertainty in practice. Illustrative examples with synthetic and experimental data are presented in Section 6, where the proposed method is compared with conventional GP model. It is also applied to SHM of two tall buildings where the dynamic properties exhibiting amplitude dependence in natural frequency and damping ratio during a typhoon event is investigated.

2. Problem context

Conventionally, data driven models assume that the training data are known precisely without uncertainty. Fig. 1 shows the schematic diagram for this situation. The set of hyper parameters ψ of the data driven model are inferred directly based
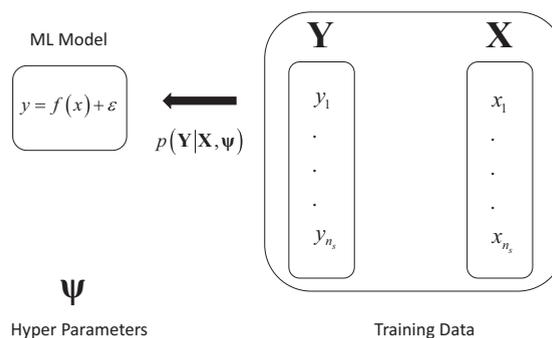

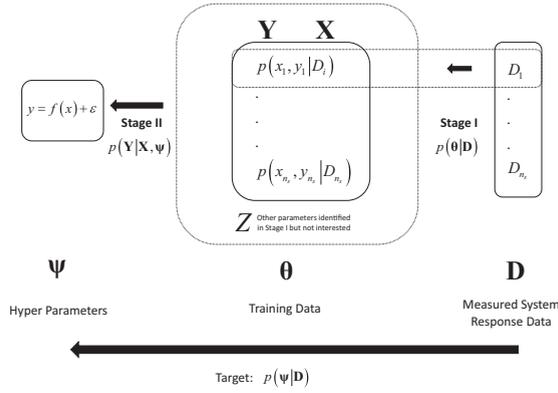**Fig. 1.** Conventional data driven model.

Fig. 2. Proposed framework.

on the input and output training sets. This is not always the case in applications, however. Both the input and output training data may not be directly measured but are rather identified based on observations or response measurements from the system; and different training points may have different precision arising from different identification uncertainties.

In SHM for example, the relationship between the modal parameters (e.g., natural frequencies, damping ratios and mode shapes) of the structures and environmental/operational conditions (e.g., temperature and power spectral density of the excitation) is often investigated using data-driven models. In this context, both the dynamic properties and environmental/-operational conditions may not be directly obtained but are identified based on the measured vibration response (e.g., acceleration data). The quality of the structural response data may vary due to sensor noise and measurement duration etc., which leads to varying levels of uncertainty in the derived training data. When the identified dynamic properties and environmental/operational conditions are used as training data to infer the hyper parameters of the data driven model, their identification uncertainty should be taken in consideration so that the quality of the training data is properly accounted for. This work focuses on such context. Fig. 2 shows the schematic diagram of the problem described above. The input and output training data $\{x_i, y_i\}$ ($i = 1, ..., n_s$ where n_s is the number of training points) are both identified from system measurements D_i. Acknowledging limited data and imperfect model, the 'exact' value of the quantity used as training data is unknown, or philosophically speaking, does not exist. Only the posterior distribution of the input and output data given the measurement **D** and in the context of identification model, i.e., $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ is available.

Making a probabilistic prediction of y for a given x requires the posterior distribution of hyper parameters ψ given the measurement **D**, i.e., $p(\psi|\mathbf{D})$. This distribution is analytically intractable in general, however. As one key theoretical contribution in this work, we will express $p(\psi|\mathbf{D})$, in terms of the posterior distribution of the training data, i.e., $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$, and the conventional marginal distribution for data driven models, i.e., $p(\mathbf{Y}|\mathbf{X}, \psi)$, in order to facilitate computation.

3. Bayesian framework

Let θ be a set of parameters identified from the available measured data set **D**. They contain three groups:

$$\theta = [\mathbf{X}, \mathbf{Y}, \mathbf{Z}] \tag{1}$$

Here, **X** and **Y** are used as input and output training data for inferring the data driven model, respectively; **Z** contains the remaining parameters identified from **D** but not related to the data driven model. Let ψ denote the hyper parameters that describe the data driven model that gives a probabilistic description of **Y** given **X**. Using the theorem of total probability, the marginal distribution of ψ given **D** is

$$p(\psi|\mathbf{D}) = \int p(\psi, \theta|\mathbf{D})d\theta \tag{2}$$

Using Baye's theorem,

$$p(\psi, \theta|\mathbf{D}) = \frac{p(\mathbf{D}|\psi, \theta)p(\psi, \theta)}{p(\mathbf{D})} \tag{3}$$

Substituting Eq. (3) into Eq. (2) gives

$$p(\psi|\mathbf{D}) = p(\mathbf{D})^{-1} \int p(\mathbf{D}|\psi, \theta)p(\psi, \theta)d\theta \tag{4}$$

Given θ, the probability distribution of **D** can be fully determined via $p(\mathbf{D}|\theta)$. The additional information from ψ is therefore redundant, i.e.,

$$p(\mathbf{D}|\boldsymbol{\psi},\boldsymbol{\theta}) = p(\mathbf{D}|\boldsymbol{\theta}) \tag{5}$$

Further applying Baye's theorem gives,

$$p(\mathbf{D}|\boldsymbol{\theta}) = \frac{p(\boldsymbol{\theta}|\mathbf{D})p(\mathbf{D})}{p(\boldsymbol{\theta})} \tag{6}$$

Substituting Eqs. (5) and (6) into Eq. (4) gives

$$p(\boldsymbol{\psi}|\mathbf{D}) = \int p(\boldsymbol{\theta}|\mathbf{D})p(\boldsymbol{\psi},\boldsymbol{\theta})p(\boldsymbol{\theta})^{-1}d\boldsymbol{\theta} \tag{7}$$

Eq. (7) is now expressed in terms of the posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{D}$, i.e. $p(\boldsymbol{\theta}|\mathbf{D})$, which encapsulates the posterior uncertainty of $\boldsymbol{\theta}$. However, the equation still contains information about $\mathbf{Z}$ (inside $\boldsymbol{\theta}$), which is redundant when making inference about $\boldsymbol{\psi}$. It is also necessary to rewrite $p(\boldsymbol{\psi},\boldsymbol{\theta})$ in a more tractable form.

Recall $\boldsymbol{\theta} = [\mathbf{X},\mathbf{Y},\mathbf{Z}]$ from Eq. (1),

$$p(\boldsymbol{\psi},\boldsymbol{\theta}) = p(\mathbf{X},\mathbf{Y},\mathbf{Z},\boldsymbol{\psi}) = p(\mathbf{X},\mathbf{Y},\mathbf{Z}|\boldsymbol{\psi})p(\boldsymbol{\psi}) \tag{8}$$

Note that $\mathbf{Z}$ is not related to $\boldsymbol{\psi}$. Assuming that $\{\mathbf{X},\mathbf{Y}\}$ and $\mathbf{Z}$ are conditionally independent for a given $\boldsymbol{\psi}$:

$$p(\mathbf{X},\mathbf{Y},\mathbf{Z}|\boldsymbol{\psi}) = p(\mathbf{X},\mathbf{Y}|\boldsymbol{\psi})p(\mathbf{Z}) \tag{9}$$

Substituting Eq. (9) into Eq. (8) gives

$$p(\boldsymbol{\psi},\boldsymbol{\theta}) = p(\mathbf{X},\mathbf{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})p(\mathbf{Z}) \tag{10}$$

Eq. (9) also implies that $\{\mathbf{X},\mathbf{Y}\}$ and $\mathbf{Z}$ are unconditionally independent:

$$
\begin{aligned}
p(\mathbf{X},\mathbf{Y},\mathbf{Z}) &= \int p(\mathbf{X},\mathbf{Y},\mathbf{Z}|\boldsymbol{\psi})p(\boldsymbol{\psi})d\boldsymbol{\psi} \\
&= p(\mathbf{Z})\int p(\mathbf{X},\mathbf{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})d\boldsymbol{\psi} \\
&= p(\mathbf{X},\mathbf{Y})p(\mathbf{Z})
\end{aligned}
\tag{11}
$$

Substituting Eqs. (10) and (11) (noting that $p(\boldsymbol{\theta}) = p(\mathbf{X},\mathbf{Y},\mathbf{Z})$) into Eq. (7), the posterior distribution can now be expressed as:

$$
\begin{aligned}
p(\boldsymbol{\psi}|\mathbf{D}) &= \int\int\int p(\mathbf{X},\mathbf{Y},\mathbf{Z}|\mathbf{D})d\mathbf{Z}\, p(\mathbf{X},\mathbf{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})p(\mathbf{X},\mathbf{Y})^{-1}d\mathbf{X}d\mathbf{Y} \\
&= \int\int p(\mathbf{X},\mathbf{Y}|\mathbf{D})p(\mathbf{X},\mathbf{Y}|\boldsymbol{\psi})p(\boldsymbol{\psi})p(\mathbf{X},\mathbf{Y})^{-1}d\mathbf{X}d\mathbf{Y} \\
&= \int\int p(\mathbf{X},\mathbf{Y}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})p(\mathbf{X},\boldsymbol{\psi})p(\mathbf{X},\mathbf{Y})^{-1}d\mathbf{X}d\mathbf{Y} \\
&= \int\int p(\mathbf{X},\mathbf{Y}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})p(\boldsymbol{\psi}|\mathbf{X})p(\mathbf{Y}|\mathbf{X})^{-1}d\mathbf{X}d\mathbf{Y}
\end{aligned}
\tag{12}
$$

It is reasonable to assume that $p(\boldsymbol{\psi}|\mathbf{X})$ is slow-varying with respect to $\boldsymbol{\psi}$ compared to $p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})$ since only knowing $\mathbf{X}$ does not provide much information about $\boldsymbol{\psi}$ due to the absence of $\mathbf{Y}$. On the other hand, in the absence of knowledge about $\boldsymbol{\psi}$ that characterises the probabilistic description of $\mathbf{Y}$ given $\mathbf{X}$, $p(\mathbf{Y}|\mathbf{X})$ is slow-varying with respect to $\mathbf{X}$ and $\mathbf{Y}$ compared to $p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})$ and hence can be assumed practically constant. We can now express $p(\boldsymbol{\psi}|\mathbf{D})$ as

$$p(\boldsymbol{\psi}|\mathbf{D}) \propto \int\int p(\mathbf{X},\mathbf{Y}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})d\mathbf{X}d\mathbf{Y} \tag{13}$$

Eq. (13) provides the means for incorporating the identification uncertainty of training data when inferring about the hyper parameters of the data driven model. It expresses the posterior distribution of $\boldsymbol{\psi}$ in terms of two posterior probability density functions (PDFs), i.e., the posterior PDF of training set given the system measurements and the posterior PDF of output data given input training data and hyperparameters. Compared to Eq. (2), Eq. (13) is computationally tractable since the first term $p(\mathbf{X},\mathbf{Y}|\mathbf{D})$ results directly from Bayesian inference of $\{\mathbf{X},\mathbf{Y}\}$ based on measurement $\mathbf{D}$ and the second term $p(\mathbf{Y}|\mathbf{X},\boldsymbol{\psi})$ results directly from the data driven model adopted.

4. Application to OMA data using GP model

The Bayesian framework in the previous section is generally applicable as long as the assumptions are met. In this section, it is specialised to OMA context where $\mathbf{D}$ is a set of ambient vibration measurements. The training data is what one would like to explain a potential phenomenon, e.g., amplitude dependence of natural frequency and damping in structural-wind engineering. In this context, the input training $\mathbf{X}$ can be modal force PSD or other environmental conditions and the output training data $\mathbf{Y}$ can comprise modal properties of interest such as natural frequencies or damping ratios. Both $\mathbf{X}$ and $\mathbf{Y}$ with their associated uncertainties may be identified based on $\mathbf{D}$ using OMA techniques. The target is to construct a data driven model for predicting modal properties $\mathbf{Y}$ as a function of environmental or operational variations $\mathbf{X}$. A Gaussian Process model is adopted for the relationship between $\mathbf{X}$ and $\mathbf{Y}$.

Recall that Eq. (13) is expressed in terms of two posterior PDFs. In Section 4.1, the first PDF $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ will be derived following an existing Bayesian OMA approach. The second PDF $p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\psi})$ will be derived based on Gaussian GP model in Section 4.2. An efficient formulation is developed in Section 4.3 by considering the characteristics of posterior uncertainty for OMA.

4.1. Bayesian operational modal analysis

First consider $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ in the context of OMA. For given measurement set D_i, the posterior distribution $p(x_i, y_i|D_i)$ can be obtained using a Bayesian operational modal analysis (BAYOMA) approach. Compared to non-Bayesian approaches, Bayesian approach views modal identification as an inference problem. The modal parameters are encapsulated in the posterior distribution given measured data and modelling assumptions, which fundamentally addresses the identification uncertainty through its posterior covariance. The identification formulations include using time domain data [31], frequency domain data based on sample power spectral density [32–34] and Fast Fourier Transform (FFT) [35,36]. Approaches considering different situations in real implementation, e.g., asynchronous data [37,38] and buried mode [39], have also been developed recently. Applications of BAYOMA approaches in SHM of modern civil infrastructures include super tall buildings [40,41] and long-span bridges [42,43].

A frequency domain Bayesian OMA method [17,44] based on FFT of ambient vibration measurements is adopted in this work. It does not involve averaging concept (not so for power spectral density based methods) and allow one to make inference based on the FFT data within a selected band around the mode(s) of interested (not so for time domain methods), which simplifies the identification model and reduces the modelling error. The main theory is reviewed in this section. The data D_i is now the FFT of ambient acceleration measurement within a selected frequency band containing the modes of interest, say $\{F_k\}$, and $\{x_i, y_i\}$ are the parameters in $\boldsymbol{\theta}$ to be identified from D_i. At frequency f_k, the FFT F_k is modelled as

$$F_k = \sum_{j=1}^{n_m} \boldsymbol{\varphi}_j \ddot{\eta}_{jk} + \boldsymbol{\varepsilon}_k \tag{14}$$

where $\boldsymbol{\varphi}_j$ and $\ddot{\eta}_{jk}$ denote the mode shape and FFT of modal acceleration of the j th contributing mode in the band ($j = 1, ..., n_m$ where n_m is the total number of contributing modes in the frequency band of interest), respectively; $\boldsymbol{\varepsilon}_k$ is the scaled FFT of prediction error arising from measurement noise or modelling error. Assuming classically damped modes, the modal acceleration satisfies the modal equation of motion:

$$\ddot{\eta}_j(t) + 2\zeta_j \omega_j \dot{\eta}_j(t) + \omega_j^2 \eta_j(t) = p_j(t) \tag{15}$$

where $\omega_j = 2\pi f_j \text{(rad/s)}$; $f_j \text{(Hz)}$, ζ_j and $p_j(t)$ are the natural frequency, damping ratio and modal force of the mode, respectively.

Taking FFT on both sides of Eq. (15) and re-arranging (noting that $\eta_{jk} = -\ddot{\eta}_{jk}/\omega_k^2$ and $\dot{\eta}_{ik} = -\mathbf{i}\ddot{\eta}_{jk}/\omega_k$) gives

$$\ddot{\eta}_{jk} = \left[1 - \beta_{jk}^2 - \mathbf{i}(2\zeta\beta_{jk}) \right]^{-1} p_{jk}, \quad \beta_{jk} = f_j/f_k \tag{16}$$

The unknown modal forces are modelled as a stationary stochastic process with a constant PSD matrix within the resonance band. The prediction error is modelled as Gaussian band-limited white noise, independent among different measurement channels and with a common PSD constant within the resonance band. Substituting Eq. (16) into Eq. (14), the theoretical PSD matrix of data is then given by

$$\mathbf{E}_k = E[F_k F_k^*] = \boldsymbol{\Phi} \mathbf{H}_k \boldsymbol{\Phi}^{\mathrm{T}} + S_e \mathbf{I}_n \tag{17}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1 \cdots \boldsymbol{\varphi}_{n_m}] \in R^{n \times n_m}$; S_e is the (constant) PSD of prediction error, $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and $\mathbf{H}_k \in C^{n_m \times n_m}$ is the theoretical PSD matrix of modal acceleration given by:

$$\mathbf{H}_k = diag(\mathbf{h}_k)\mathbf{S}diag(\mathbf{h}_k^*) \tag{18}$$

Here, $\mathbf{S}$ is the PSD matrix of modal forces, $\mathbf{h}_k \in C^{n_m}$ is the vector of modal frequency response function with

$$h_{jk} = \left[1 - \beta_{jk}^2 - \mathbf{i}(2\zeta\beta_{jk}) \right]^{-1} \tag{19}$$

and $diag(\mathbf{h}_k)$ denotes a diagonal matrix with the j th element equal to h_{jk}.

Using Bayes' theorem and assuming a uniform prior distribution, the PDF of $\boldsymbol{\theta}$ given $\{F_k\}$ is proportional to the likelihood function, i.e.,

$$p(\boldsymbol{\theta}|\{F_k\}) \propto p(\{F_k\}|\boldsymbol{\theta}) \tag{20}$$

Assuming long data duration, $\{F_k\}$ are asymptotically independent at different frequencies and jointly (circularly symmetric) complex Gaussian [45]. The likelihood function is then given by:

$$p(\{F_k\}|\boldsymbol{\theta}) = (\pi)^{-nN_f} \times \prod_k (\det \mathbf{E}_k)^{-1} \exp\left[-\sum_k F_k^* \mathbf{E}_k^{-1} F_k\right] \tag{21}$$

It is more convenient to write:

$$p(\boldsymbol{\theta}|\{F_k\}) \propto \exp[-L(\boldsymbol{\theta})] \tag{22}$$

where

$$L(\boldsymbol{\theta}) = \sum_k \ln\det\mathbf{E}_k + \sum_k F_k^* \mathbf{E}_k^{-1} F_k \tag{23}$$

is the 'negative log-likelihood function' (NLLF).

For sufficient data, modal identification problem is 'globally identifiable' [46]. The posterior PDF (i.e., $p(\boldsymbol{\theta}|\{F_k\})$) then has a centralised shape with a unique peak at the most probable value (MPV). The MPV of $\boldsymbol{\theta}$ can be determined by maximising the posterior PDF (or equivalently minimising the NLLF with respect to $\boldsymbol{\theta}$). The posterior PDF can be approximated by a Gaussian PDF [46],

$$p(\boldsymbol{\theta}|\{F_k\}) \approx (2\pi)^{-n_\theta/2} \left(\det\hat{\mathbf{C}}\right)^{-1/2} \exp\left[-\frac{1}{2}\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^T \hat{\mathbf{C}}^{-1}\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)\right] \tag{24}$$

where n_θ is the number of parameters in $\boldsymbol{\theta}$; $\hat{\boldsymbol{\theta}}$ is the MPV and $\hat{\mathbf{C}}$ is the posterior covariance matrix, equal to the inverse of Hessian of NLLF at MPV. The posterior covariance of a particular parameter in $\boldsymbol{\theta}$ can be extracted from the corresponding entry of $\hat{\mathbf{C}}$.

When the training data $\{x_i, y_i\}$ is taken from $\boldsymbol{\theta}$, for a given data D_i, $p(x_i, y_i|D_i)$ can be expressed as:

$$p(x_i, y_i|D_i) \approx (2\pi)^{-1}\left(\det\hat{\mathbf{C}}_i\right)^{-1/2} \exp\left[-\frac{1}{2}\left(\begin{bmatrix} x_i \\ y_i \end{bmatrix} - \begin{bmatrix} \hat{x}_i \\ \hat{y}_i \end{bmatrix}\right)^T \hat{\mathbf{C}}_i^{-1}\left(\begin{bmatrix} x_i \\ y_i \end{bmatrix} - \begin{bmatrix} \hat{x}_i \\ \hat{y}_i \end{bmatrix}\right)\right] \tag{25}$$

where $\hat{\mathbf{C}}_i$ is the partition of $\hat{\mathbf{C}}$ with respect to x_i and y_i.

The identification results of $\{x_i, y_i\}$ from D_i will be taken as training data for constructing the Gaussian process model in order to learn the relationship between $\mathbf{X}$ and $\mathbf{Y}$. Given $\{x_i, y_i\}_{i=1}^{n_s}$, the system measurements $\{D_i\}_{i=1}^{n_s}$ are assumed to be independent. Together with the fact that D_i only depends on $\{x_i, y_i\}$, we have $p(\mathbf{D}|\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n_s} p(D_i|\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n_s} p(D_i|x_i, y_i)$. Consequently, using Baye's theorem with a flat prior on $(\mathbf{X}, \mathbf{Y})$, we have

$$p(\mathbf{X}, \mathbf{Y}|\mathbf{D}) \propto p(\mathbf{D}|\mathbf{X}, \mathbf{Y}) = \prod_{i=1}^{n_s} p(D_i|x_i, y_i) \propto \prod_{i=1}^{n_s} p(x_i, y_i|D_i) = N\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}\middle|\begin{bmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{Y}} \end{bmatrix}, \begin{bmatrix} \mathbf{C_X} & \mathbf{C_{XY}} \\ \mathbf{C_{YX}} & \mathbf{C_Y} \end{bmatrix}\right) \tag{26}$$

which is a Gaussian PDF with $\mathbf{C_X} = diag([c_{x_1}...c_{x_{n_s}}])$, $\mathbf{C_Y} = diag\left([c_{y_1}...c_{y_{n_s}}]\right)$ and $\mathbf{C_{XY}} = \mathbf{C_{YX}} = diag\left([c_{x_1 y_1}...c_{x_{n_s} y_{n_s}}]\right)$.

4.2. Gaussian process model

Now consider $p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\psi})$ for a GP model. Without loss of generality, a regression model with unknown relationship f between a given input x and output y can be written as:

$$y = f(x) + e \tag{27}$$

where e accounts for modelling error. Instead of parameterising f, a GP model assumes that given the input data $\mathbf{X}$ the output data $\mathbf{Y}$ are jointly Gaussian:

$$\mathbf{Y} \sim GP(\mathbf{M}, \mathbf{K} + \sigma_e^2 \mathbf{I}) \tag{28}$$

It is characterised by the mean $\mathbf{M}$ and covariance $\mathbf{K}$, which are functions of the input training data $\mathbf{X}$ and hyperparameters $\boldsymbol{\psi}$. The modelling error e is assumed to be Gaussian and its variance σ_e^2 is a hyper parameter as well. Accordingly,

$$p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\psi}) = (2\pi)^{-n_s/2}\det(\mathbf{K} + \sigma_e^2 \mathbf{I})^{-1/2}\exp\left(-\frac{1}{2}(\mathbf{Y} - \mathbf{M})^T(\mathbf{K} + \sigma_e^2 \mathbf{I})^{-1}(\mathbf{Y} - \mathbf{M})\right) \tag{29}$$

4.3. Efficient formulation with BAYOMA

As apparent in Eq. (13), evaluating $p(\boldsymbol{\psi}|\mathbf{D})$ requires integrating the product of PDFs with respect to both input and output training data. This may not be analytically tractable. For general cases, advanced numerical tools such as Monte Carlo inte-

gration may be pursued. Here we develop a strategy that takes advantage of the properties of $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ for OMA to facilitate computation.

First rewrite $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ as

$$p(\mathbf{X}, \mathbf{Y}|\mathbf{D}) = p(\mathbf{Y}|\mathbf{X}, \mathbf{D})p(\mathbf{X}|\mathbf{D}) \tag{30}$$

For OMA, $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$ is a Gaussian PDF (see Eq. (26)). Clearly, $p(\mathbf{X}|\mathbf{D})$ is a Gaussian PDF with mean $\hat{\mathbf{X}}$ and covariance matrix $\mathbf{C_X}$. On the other hand, $p(\mathbf{Y}|\mathbf{X}, \mathbf{D})$ is the conditional PDF, which from standard results is also a Gaussian PDF for $\mathbf{Y}$ with mean $\hat{\mathbf{Y}} + \mathbf{C_{YX}}\mathbf{C_X}^{-1}(\mathbf{X} - \hat{\mathbf{X}})$ and covariance matrix $\mathbf{C_Y} - \mathbf{C_{YX}}\mathbf{C_X}^{-1}\mathbf{C_{XY}}$.

Substituting Eq. (30) into Eq. (13) and swapping the sequence of integration gives

$$p(\boldsymbol{\psi}|\mathbf{D}) \propto \int \underbrace{\int p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\psi})p(\mathbf{Y}|\mathbf{X}, \mathbf{D})d\mathbf{Y}}_{F(\mathbf{X})} \; p(\mathbf{X}|\mathbf{D})d\mathbf{X} \tag{31}$$

The inner integrand $p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\psi})p(\mathbf{Y}|\mathbf{X}, \mathbf{D})$ is a product of two Gaussian PDFs with respect to $\mathbf{Y}$, which can be integrated analytically. The result is a Gaussian-like form:

$$F(\mathbf{X}) = (2\pi)^{-n_s/2}\det(\mathbf{C_W})^{-1/2}\exp\left(-\frac{1}{2}\mathbf{W}^T\mathbf{C_W}^{-1}\mathbf{W}\right) \tag{32}$$

where

$$\mathbf{W} = \hat{\mathbf{Y}} + \mathbf{C_{YX}}\mathbf{C_X}^{-1}(\mathbf{X} - \hat{\mathbf{X}}) - \mathbf{M} \tag{33}$$

$$\mathbf{C_W} = \mathbf{C_Y} - \mathbf{C_{YX}}\mathbf{C_X}^{-1}\mathbf{C_{XY}} + \mathbf{K} + \sigma_e^2\mathbf{I} \tag{34}$$

It is interesting to note that $\mathbf{W}$ is simply the difference of the mean vectors of these two Gaussian PDFs; and $\mathbf{C_W}$ is the sum of the covariance matrices.

The resulting integrand $F(\mathbf{X})p(\mathbf{X}|\mathbf{D})$ in Eq. (31) generally depends on $\mathbf{X}$ in a nonlinear manner and is not proportional to a standard distribution. Without resorting to brute-force numerical integration that is prohibitive, a Gaussian type approximation as in Section 3.2.2 of [47] is adopted, which gives

$$p(\boldsymbol{\psi}|\mathbf{D}) \propto \int F(\mathbf{X})p(\mathbf{X}|\mathbf{D})d\mathbf{X} \approx (2\pi)^{-n_s/2}\det(\mathbf{C'_W})^{-1/2}\exp\left(-\frac{1}{2}\mathbf{W}'^T\mathbf{C'_W}^{-1}\mathbf{W}'\right) \tag{35}$$

where

$$\begin{aligned}\mathbf{W}' &= \int \mathbf{W}N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} \\ &= \hat{\mathbf{Y}} - \mathbf{M}'\end{aligned} \tag{36}$$

$$\begin{aligned}\mathbf{C'_W} = \quad & \mathbf{C_Y} + \sigma_e^2\mathbf{I} + \mathbf{K}' \\ - \quad & \mathbf{C_{XY}}\mathbf{C_X}^{-1}\int\left(\mathbf{X} - \hat{\mathbf{X}}\right)\mathbf{M}^T N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} \\ - \quad & \mathbf{C_{XY}}\mathbf{C_X}^{-1}\int\mathbf{M}\left(\mathbf{X} - \hat{\mathbf{X}}\right)^T N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} \\ + \quad & \int\mathbf{M}\mathbf{M}^T N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} - \mathbf{M}'\mathbf{M}'^T\end{aligned} \tag{37}$$

with

$$\mathbf{M}' = \int \mathbf{M}N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} \tag{38}$$

$$\mathbf{K}' = \int \mathbf{K}N\left(\mathbf{X}\big|\hat{\mathbf{X}}, \mathbf{C_X}\right)d\mathbf{X} \tag{39}$$

A detailed investigation of the quality of this approximation can be found in [47]. The approximation is briefly explained in the appendix. Whether the analytical expressions of $\mathbf{W}'$ and $\mathbf{C'_W}$ (i.e., Eq. (36) and (37)) are available still depends on the form of the mean and covariance function. For a zero mean function (i.e., $\mathbf{M} = \mathbf{0}$, commonly assumed for GP models), $\mathbf{W}'$ and $\mathbf{C'_W}$ can be simplified as:

$$\mathbf{W}' = \hat{\mathbf{Y}} \tag{40}$$

$$\mathbf{C'_W} = \mathbf{C_Y} + \sigma_e^2\mathbf{I} + \mathbf{K'} \tag{41}$$

which no longer depend on the posterior correlation $\mathbf{C_{XY}}$. On the other hand, it turns out [47] that when the covariance function $\mathbf{K}$ is linear, Gaussian, polynomial or a mixture of these three, $\mathbf{K'}$ can be evaluated analytically. Consider the squared exponential covariance function (which is widely used), i.e.,

$$\mathbf{K}(x_i, x_j) = \sigma_f^2\exp\left(-\frac{1}{2}(x_i - x_j)^T w^{-1}(x_i - x_j)\right) \tag{42}$$

where σ_f and w are the hyper parameters that denote the signal variance and characteristic length-scale, the analytical expression of $\mathbf{K'}$ can be obtained as [28]

$$\mathbf{K'}(x_i, x_j) = \frac{\sigma_f^2\exp\left(-\frac{1}{2}(\hat{x}_i - \hat{x}_j)^T\left(w + c_{x_i} + c_{x_j}\right)^{-1}(\hat{x}_i - \hat{x}_j)\right)}{\left|1 + w^{-1}\left(c_{x_i} + c_{x_j}\right)\left(1 - \delta_{ij}\right)\right|^{1/2}} \tag{43}$$

where δ_{ij} is the Kronecker delta, i.e., $\delta_{ij} = 1$ if $i = j$ and zero otherwise.

For inferring the hyper parameters ψ, it is more convenient to work with the negative log-likelihood function

$$L_\psi = \frac{1}{2}\mathrm{lndet}(\mathbf{C'_W}) + \frac{1}{2}\mathbf{W'}^T\mathbf{C'_W}^{-1}\mathbf{W'} \tag{44}$$

such that

$$p(\psi|\mathbf{D}) \propto \exp(-L_\psi) \tag{45}$$

The hyper parameter ψ now can be obtained by maximising $p(\psi|\mathbf{D})$, or equivalently minimising L_ψ.

5. Further simplifications in specialised OMA

Eq. (44) applies to the general condition of OMA with long data. Below we consider some special situations that further take advantage of characteristics of identification uncertainty in OMA.

5.1. Neglecting posterior correlation

For well separated modes, long data duration and small damping (which is commonly the case in OMA), the posterior correlation between any pair among the identified modal parameters are asymptotically small (except for the posterior correlation between damping ratio and modal force PSD) [48,49]. When $\{x_i, y_i\}$ come from such uncorrelated pairs, Eq. (26) can be expressed as:

$$\begin{aligned}
p(\mathbf{X}, \mathbf{Y}|\mathbf{D}) &= \prod_{i=1}^{n_s} p(x_i|D_i)p(y_i|D_i) \\
&= p(\mathbf{X}|\mathbf{D})p(\mathbf{Y}|\mathbf{D}) \\
&= N\left(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\right)N\left(\mathbf{Y}\big|\hat{\mathbf{Y}},\mathbf{C_Y}\right)
\end{aligned} \tag{46}$$

Substituting Eq. (46) into Eq. (13) gives

$$p(\psi|\mathbf{D}) \propto \int\int p(\mathbf{Y}|\mathbf{D})p(\mathbf{X}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\psi)d\mathbf{X}d\mathbf{Y} = \int\int p(\mathbf{X}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\psi)d\mathbf{X}p(\mathbf{Y}|\mathbf{D})d\mathbf{Y} \tag{47}$$

The integrand $p(\mathbf{X}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\psi)$ depends on $\mathbf{X}$ in a non-linear manner. Applying an approximation similar to Eq. (35) gives

$$\int p(\mathbf{X}|\mathbf{D})p(\mathbf{Y}|\mathbf{X},\psi)d\mathbf{X} \approx N(\mathbf{Y}|\mathbf{M'},\mathbf{K''} + \sigma_e^2\mathbf{I}) \tag{48}$$

where

$$\mathbf{M'} = \int \mathbf{M}N\left(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} \tag{49}$$

$$\mathbf{K''} = \mathbf{K'} + \int \mathbf{M}\mathbf{M}^TN\left(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} - \mathbf{M'}\mathbf{M'}^T \tag{50}$$

Substituting Eq. (48) into Eq. (47) gives a product of two Gaussian PDFs with respect to $\mathbf{Y}$, which yields a scaled Gaussian PDF. Integrating $\mathbf{Y}$ out gives the scaling factor only. The resulting negative loglikelihood function now can be written as

$$L_\psi = \frac{1}{2}\text{lndet}\left(\mathbf{C_Y} + \mathbf{K}'' + \sigma_e^2\mathbf{I}\right) + \frac{1}{2}\left(\hat{\mathbf{Y}} - \mathbf{M}'\right)^T\left(\mathbf{C_Y} + \mathbf{K}'' + \sigma_e^2\mathbf{I}\right)^{-1}\left(\hat{\mathbf{Y}} - \mathbf{M}'\right) \tag{51}$$

It is easy to check that this is consistent with Eq. (44) when $\mathbf{C_{YX}} = \mathbf{C_{XY}} = \mathbf{0}$.

5.2. Neglecting uncertainty of input training data

Consider the case where the uncertainty in input training data $\mathbf{X}$ can be neglected. For example, the environmental/operational variables may be directly measured without identification process, e.g., the temperature. For such training data, the posterior PDF of input data can be practically taken as a delta function at the MPV $\hat{\mathbf{X}}$, i.e.,

$$p(\mathbf{X}|\mathbf{D}) = \delta\left(\mathbf{X} - \hat{\mathbf{X}}\right) \tag{52}$$

Substituting Eq. (52) into Eq. (31) gives

$$\begin{aligned}
p(\psi|\mathbf{D}) &\propto \int\int p(\mathbf{Y}|\mathbf{X},\mathbf{D})p(\mathbf{Y}|\mathbf{X},\psi)\delta\left(\mathbf{X} - \hat{\mathbf{X}}\right)d\mathbf{X}d\mathbf{Y} \\
&= \int p\left(\mathbf{Y}\big|\hat{\mathbf{X}},\mathbf{D}\right)p\left(\mathbf{Y}\big|\hat{\mathbf{X}},\psi\right)d\mathbf{Y} \\
&= \int N\left(\mathbf{Y}\big|\hat{\mathbf{Y}},\mathbf{C_Y}\right)N\left(\mathbf{Y}\big|\hat{\mathbf{M}},\hat{\mathbf{K}} + \sigma_e^2\mathbf{I}\right)d\mathbf{Y} \\
&= (2\pi)^{-n_s/2}\det\left(\mathbf{C_Y} + \hat{\mathbf{K}} + \sigma_e^2\mathbf{I}\right)^{-1/2}\exp\left(-\frac{1}{2}\left(\hat{\mathbf{Y}} - \hat{\mathbf{M}}\right)^T\left(\mathbf{C_Y} + \hat{\mathbf{K}} + \sigma_e^2\mathbf{I}\right)^{-1}\left(\hat{\mathbf{Y}} - \hat{\mathbf{M}}\right)\right)
\end{aligned} \tag{53}$$

where $\hat{\mathbf{M}}$ and $\hat{\mathbf{K}}$ are the conventional mean and covariance function calculated based on $\hat{\mathbf{X}}$. The corresponding negative log-likelihood function is now given by

$$\begin{aligned}
L_\psi &= \frac{n_s}{2}\ln 2\pi + \frac{1}{2}\text{lndet}\left(\mathbf{C_Y} + \hat{\mathbf{K}} + \sigma_e^2\mathbf{I}\right) \\
&\quad + \frac{1}{2}\left(\hat{\mathbf{Y}} - \hat{\mathbf{M}}\right)^T\left(\mathbf{C_Y} + \hat{\mathbf{K}} + \sigma_e^2\mathbf{I}\right)^{-1}\left(\hat{\mathbf{Y}} - \hat{\mathbf{M}}\right)
\end{aligned} \tag{54}$$

which further simplifies the computation.

Although the term $\mathbf{C_Y} + \hat{\mathbf{K}} + \sigma_e^2\mathbf{I}$ in Eq. (54) has an effect similar to simply adding non-stationary modelling errors for different training points, the nature is different. The imprecision in the training data considered in this work stems from the remaining uncertainty after using information from the ambient measurements in the context of OMA model, which is related to the quality of the measurement. The 'output uncertainty' considered in conventional data driven models refers to the probability distribution of modelling error e (see Eq. (27)). It reflects the quality of the data driven model f when expressing the output y as a function of the input x, which need not reflect consistently to the quality of measured response data.

6. Illustrative examples

Three examples are presented to compare the performance of the classic GP model (which considers the training data as precise values without uncertainty) and the proposed method that accounts for the identification uncertainty. Except for the synthetic data example, the training data are first normalised (i.e., subtracted by the sample mean and divided by sample standard deviation) to facilitate computation. The GP models used in these examples are based on the squared exponential covariance function and zero mean function. The first example is based on synthetic data, where the true function between the input and output data exists and is known in advance, so that the method can be benchmarked. The second example is based on OMA data measured from a laboratory shear building model, where physical complexity is naturally reflected in the measurements and there need not be any 'true' function between the input and output training data. Finally, the proposed method is applied to OMA data of two tall buildings during a typhoon event, which illustrates its feasibility to SHM data in real applications.

6.1. Synthetic data validation

In order to investigate the behaviour of both the conventional GP model and the GP model considering identification uncertainty, a simple sinusoidal function $f(x) = \sin x$ is considered in this example. Twenty training data points of $\mathbf{X}$ are uniformly sampled from the range $[-1, 10]$. Additional noise is added to the output of each sampling point. The posterior distribution of the output training points (i.e., $p(\mathbf{Y}|\mathbf{D})$) is assumed to be Gaussian with mean value at the output values and a posterior variance equal to the square of the added noise value. The posterior uncertainty of input training points (i.e., $p(\mathbf{X}|\mathbf{D})$) are assumed to be Gaussian distributed with mean value at the sampled input points. Three scenarios are considered.

In the first scenario, the noise is randomly generated as zero mean Gaussian with a standard deviation of 0.2. In the second scenario, five training points are randomly picked and contaminated with noise as zero mean Gaussian with a standard deviation of 1. For the remaining training points, the noise is generated as zero mean Gaussian with a standard deviation of 0.2. For the first two scenarios, the posterior variance of each input training data is set to be the same as that of the corresponding output data. The third scenario is the same as the second one except that the posterior variance of each input training data is sampled uniformly in [0, 5].

Figs. 3 and 4 show a typical set of results for the first two scenarios, respectively. The true values of the sinusoidal function are shown as a solid line. The training data are plotted as squares with error bars indicating ±one standard deviation of identification uncertainty. The conventional GP model is shown with mean values as a dotted dashed line and ±2 standard deviation (i.e., about 95% confidence interval) as the light grey area. The data driven model based on the proposed method (considering both input and output uncertainty in training data) is shown with mean values as a dashed line and ±2 standard deviation as the dark grey area. For the first scenario, the data driven models based on these two methods are basically the same, as evidenced from the variation of the mean lines in Fig. 3. This is reasonable as the identification uncertainty among the training points does not vary too much. This is not the case for the second scenario, however. Discrepancies can be found between these two methods, especially around the training points with large uncertainty. For the conventional GP method, the training points are treated equally without considering their associated uncertainties. On the other hand, the proposed method effectively put more weights on the training points with smaller uncertainty. The resulting data driven model based on the proposed method shows better behaviour compared to the conventional GP model in this case.

To investigate the effect of uncertainty in the input training data, confine now to the proposed method. Figs. 5 and 6 show the trained data driven models in the second and third scenarios, respectively. The model considering both input and output uncertainty (i.e., Eq. (44)) and the model considering output uncertainty only (i.e., Eq. (54)) are compared. As seen in Fig. 5, these two models do not differ significantly when the posterior variance of the input and output data are identical. Different trials have been conducted where posterior variance of the input data is proportional to that of the output data, which show qualitatively the same result. Discrepancies can be found between these two models when the uncertainty in the input and output training data are not systematically related (as seen in Fig. 6). Considering the uncertainty in both input and output data in this case will have a significant effect on the importance of each training point when the GP model is trained compared to the case when only the uncertainty in output training data is considered.

In this example, the identification uncertainty of the output training points is set based on the differences compared to true values of the function. It should be noted that this is not the case in real applications but is used here to benchmark the proposed method. The identification uncertainty in real application is related to the measured data from which the training point is identified. It need not be related to the 'true function' (if any) that describes the relationship between the input and output data.

6.2. Laboratory shear building model

Consider a three-storey laboratory aluminium shear building structure as shown in Fig. 7. Each floor measures 25.5cm × 30.5cm × 2.5cm and each column section measures 0.5 cm by 2.5 cm. The storey height is 8 cm for all floors. Three piezoelectric accelerometers distributed at the centre of each floor are used to measure the vibration in the weak direction (parallel to the paper).
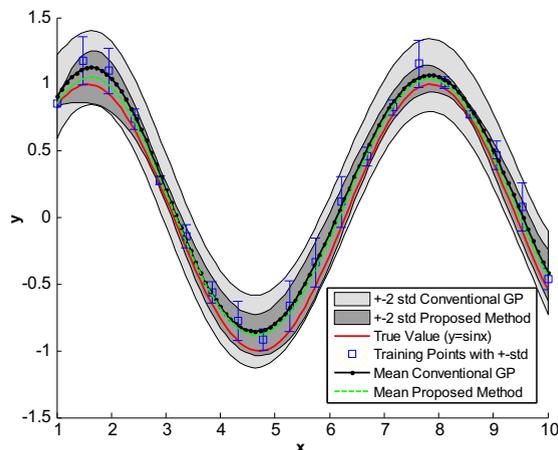

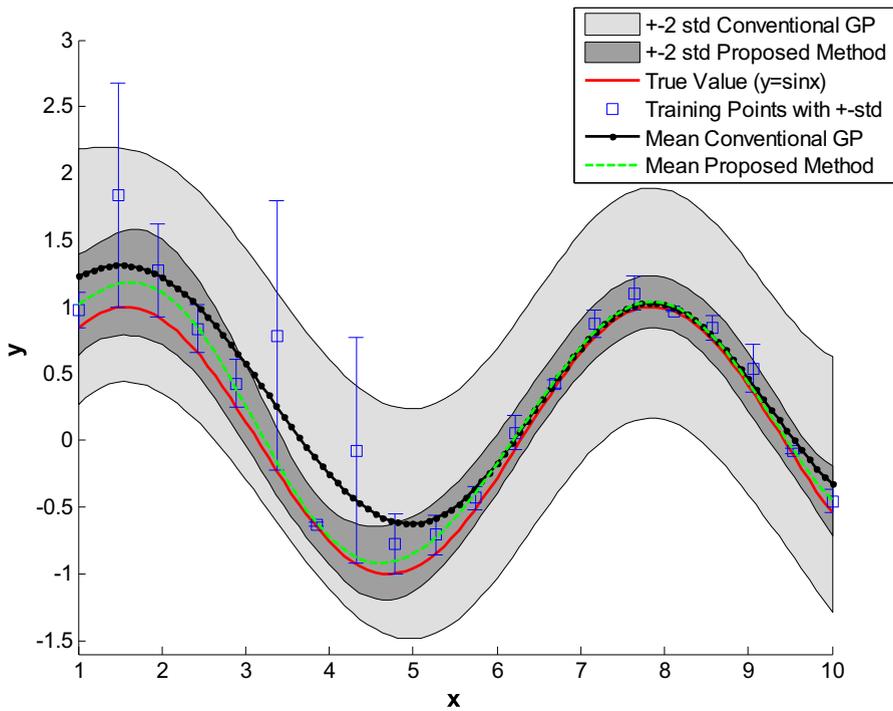**Fig. 3.** Typical learning of sinusoidal function, Scenario 1.

Fig. 4. Typical learning of sinusoidal function, Scenario 2.
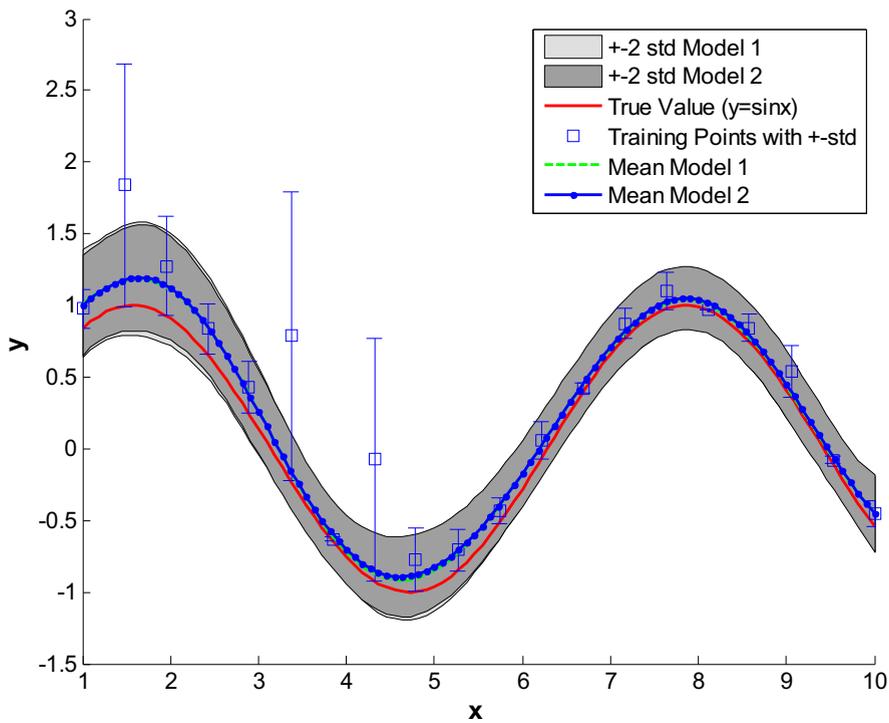

Fig. 5. Comparison between two proposed models, scenario 2 (model 1: GP model considering both x and y uncertainty; model 2: GP model considering y uncertainty only).
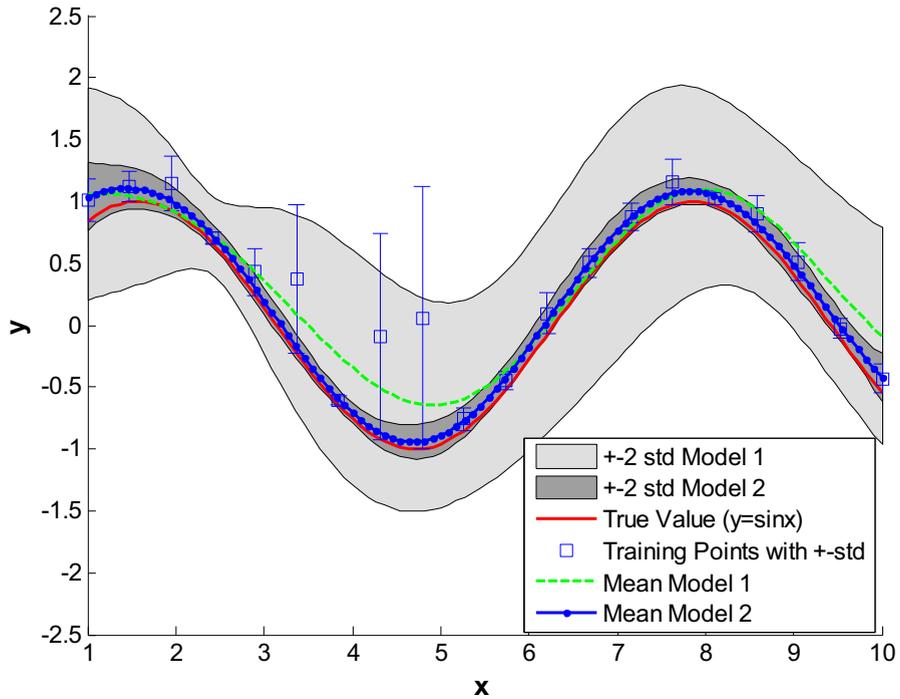
Fig. 6. Comparison between two proposed models, scenario 3 (model 1: GP model considering both x and y uncertainty; model 2: GP model considering y uncertainty only).

Fig. 7. Three-storey laboratory shear building setup.

To investigate the modal properties of the structure under different excitation levels. The structure is excited with different levels of Gaussian white noise provided by an electrodynamic shaker. The resulting modal force PSD ranges from $10^{-13} g^2/Hz$ to $10^{-6} g^2/Hz$. The vibration response of the structure was recorded with different durations (i.e., 1 min, 2 min and 5 min), which led to different uncertainty of the identified modal parameters. Fifteen sets of vibration response data were measured in total at a sampling rate of 2048 Hz. The recorded data were subsequently decimated to 512 Hz for analysis. Fig. 8 shows the root singular value (SV) spectrum of a typical data set (i.e., a plot of the square root of the eigenvalues of the real part of the spectral density matrix against frequency. This is equivalent to the singular values since the spectral density matrix is real symmetric.). Modal analysis focuses on the mode around 77 Hz. The selected frequency band is [75 80]Hz (shown as '[−]' in the figure) and the initial guess of the natural frequency is 77 Hz (shown as 'o' in the figure). Fig. 9 shows the identified mode shape based on this data set. Fig. 10 shows the identified natural frequencies of this mode against the modal force PSD among the measured data sets, where the error bar denotes ±2 standard deviation of the identification uncertainty. Different marks reflect the duration of the measured vibration data from which the modal parameters are identified. It can be seen that the identified natural frequency generally decreases with the modal force PSD.

In this example, the modal force PSD (reflecting the level of the white noise excitation generated by the shaker) is the operational parameter and is considered as the input of the data driven models. The identified natural frequencies are the modal properties of interest and considered as the output of the data driven models. The data driven models trained based
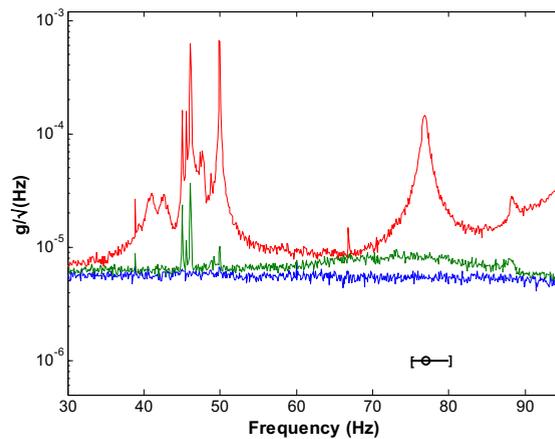

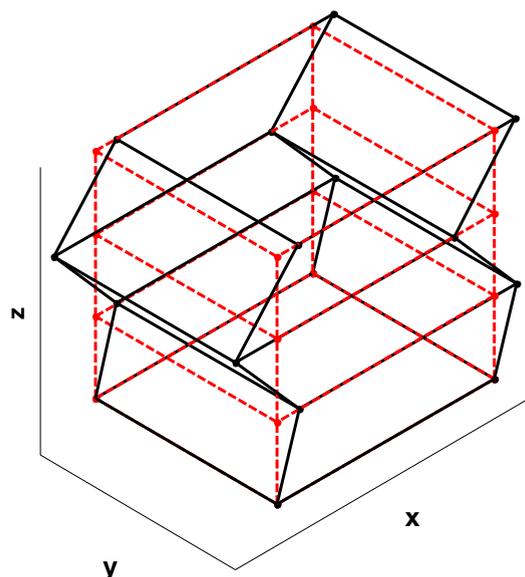
Fig. 8. Root SV spectrum, laboratory shear building.

Fig. 9. Identified mode shape, laboratory shear building (dashed line: undeformed mode shape; solid line: deformed mode shape).
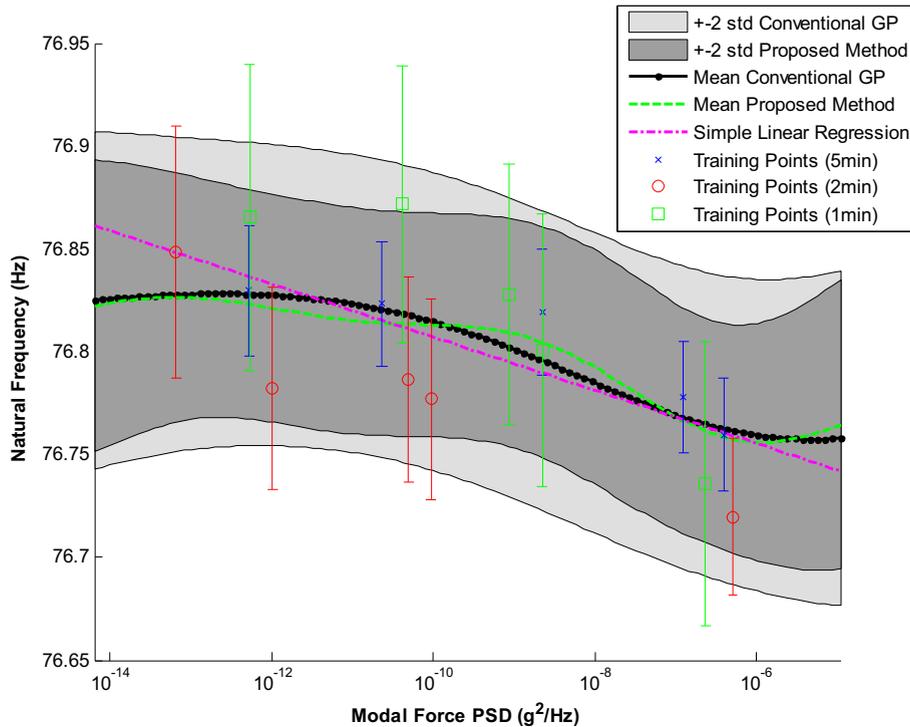
Fig. 10. Identified natural frequencies against modal force PSD with data driven models, laboratory shear building.

on the identification results are plotted in Fig. 10. Discrepancies can be observed between the conventional GP model and the proposed model that considers identification uncertainty in both the modal force PSD and natural frequency. The data driven model based on the proposed method is more sensitive to the training points with smaller uncertainty, i.e., the ones identified based on 10 min data. On the other hand, the conventional GP model simply averages out the variation of the natural frequencies with different values of identification uncertainty at different modal force PSD ranges. The results based on a simple linear regression model are also plotted in Fig. 10. Compared to the GP model, the simple linear regression model only provides a general trend without capturing the detailed relationship between the natural frequencies and the modal force PSD.

6.3. Tall building SHM under typhoon event

The field data used in this example is from a previous work [40]. Two buildings in Hongkong under a typhoon event are instrumented in this example. Building A is a tubular concrete building with a central core wall system located in the north-west of Waglan Island station in Hong Kong. It is 310 m tall and 50 m by 50 m in plan. Building B is 320 m tall and 50 m by 50 m in plan. Benchmark tests of these two buildings have been conducted under normal wind conditions with four triaxial accelerometers placed at four corners on the roof. Detailed modal identification results with mode shape plots can be found in Figs. 1 and 2 of [40].

The vibration response of the buildings was measured during Typhoon Koppu in 2009. Typhoon Koppu visited Hong Kong at about 11 am on 14th September 2009 with wind speed at Waglan ranging between 25 km/hr to 120 km/hr, which provided an opportunity to investigate the in-situ dynamic behaviour of these two buildings under strong wind. A triaxial force balance accelerometer was placed in a secure room on the roof of the buildings to measure vibration response. Forty-eight hours of acceleration time history data were recorded. The accelerometer has a noise level of $0.5\mu g/\sqrt{Hz}$ and the data was logged using a 24bit digital signal recorder at a sampling rate of 50 Hz. The whole time history data is divided into non-overlapping segments each with a duration of 30mins. The input loading and structural response are modelled as stationary stochastic process within each segment such that OMA techniques can be applied to identify the modal parameters. The investigation here focuses on applying the proposed data driven model to investigate the dynamic properties of the structure against environmental variations. Specifically, environmental variation here refers to the modal force PSD, which reflects the intensity of the wind. The natural frequency and damping ratio are the dynamic properties investigated in this example. Simple linear regression models have been applied in the previous work (see details in [40]) and the investigation here focuses on classic GP model and the proposed method.

Fig. 11 shows a typical time history data measured from Building A starting at 15th September 2009 1:35am when the wind was strong. Figs. 12 and 13 show the corresponding root PSD (i.e., the square root of power spectral density of the individual measured data channels.) and SV spectrum. Modal analysis focuses on the first two modes marked in Fig. 13, where '[−]' denotes the selected frequency band and 'o' denotes the initial guess of natural frequency. These two modes are translational modes (see mode shape plot in Fig. 14) identified simultaneously based on the same band as they are closely-spaced.
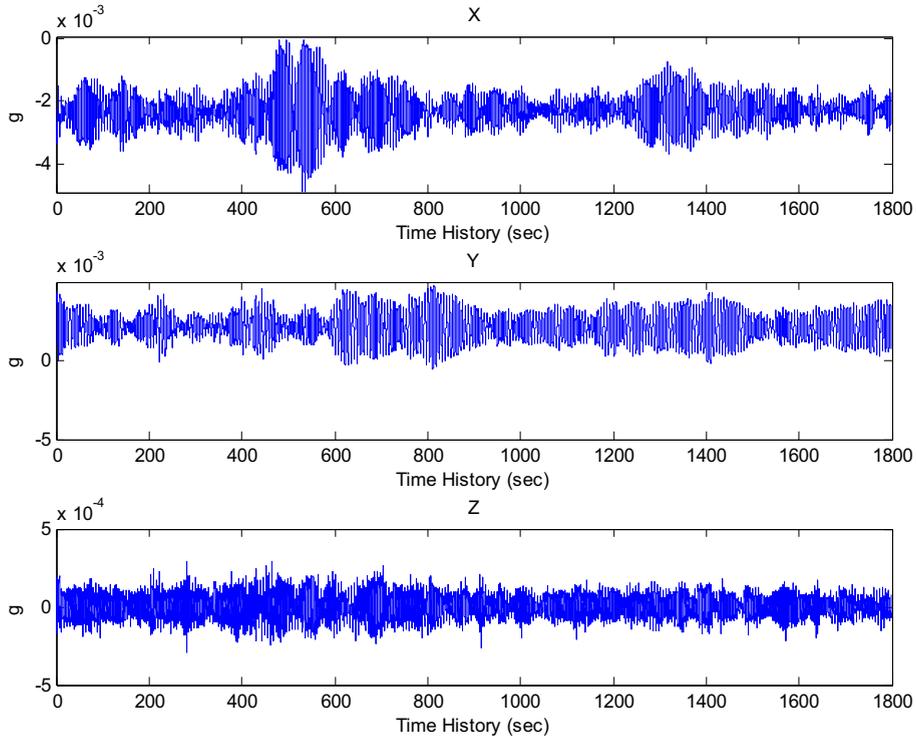
Fig. 11. Typical time history data under strong wind, building A (15th September 2009 1:35 am).
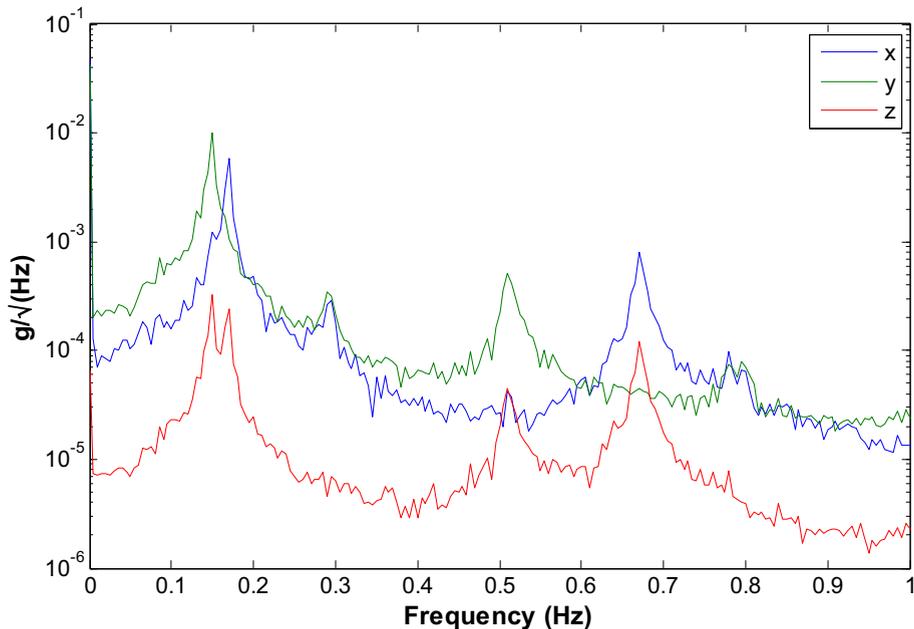
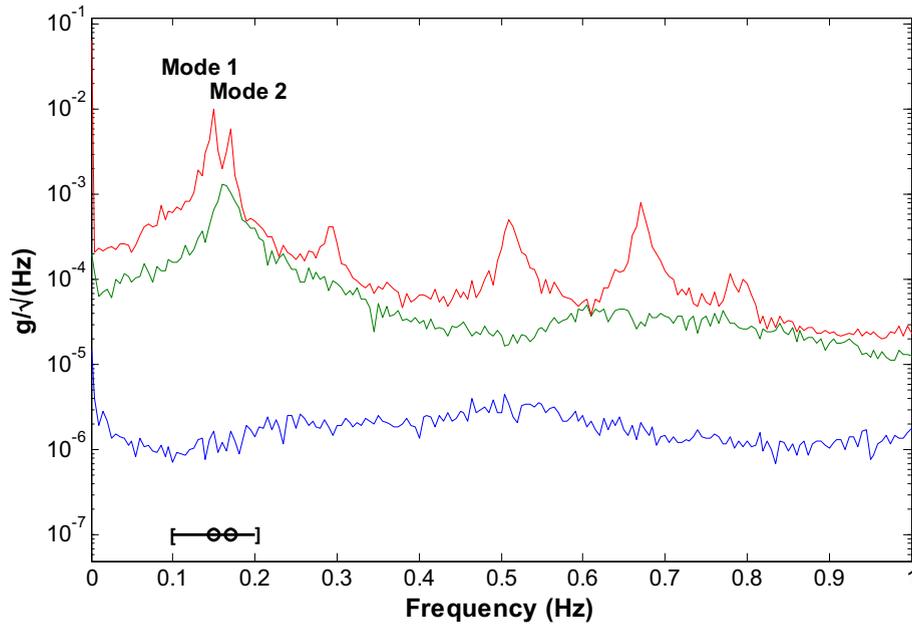**Fig. 12.** Root PSD Spectrum of Data, Building A (15th September 2009 1:35 am).

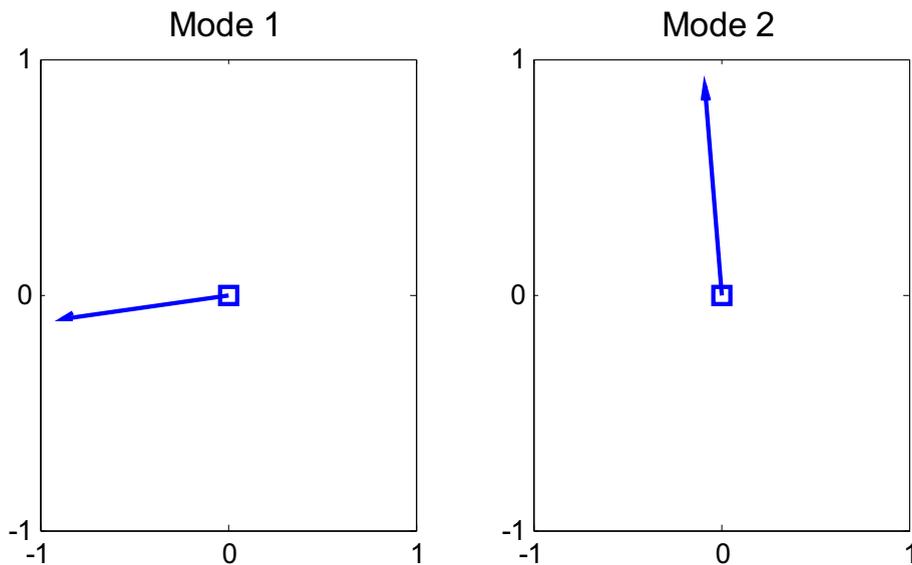**Fig. 13.** Root SV Spectrum of Data, Building A (15th September 2009 1:35 am).

Fig. 14. Identified Mode Shape, Building A (15th September 2009 1:35 am).

The interactions between these two modes also increase the identification uncertainty of the modal parameters. Figs. 15 and 16 show the identified natural frequency against the modal force PSD for Mode 1 and Mode 2, respectively. The square in the figure denotes the MPV and the error bar represents ±2 posterior standard deviation. The modal force PSD reflects the vibration amplitude of the structure. There is an inverse trend between the natural frequency and modal force PSD, indicating the amplitude dependence of the tested structure.

The proposed method and conventional GP model have been applied to investigate such amplitude dependence. The squared exponential function is selected as the covariance function and the mean function is set as zero for both models. The predictive mean values with the predictive 95% confidence bounds (i.e., ±2 standard deviation) for both methods are shown in Figs. 15 and 16 as well. It can be seen that the prediction from the proposed model is similar to that based on the conventional GP model based on the training data of mode 1. This is reasonable as the posterior uncertainties among
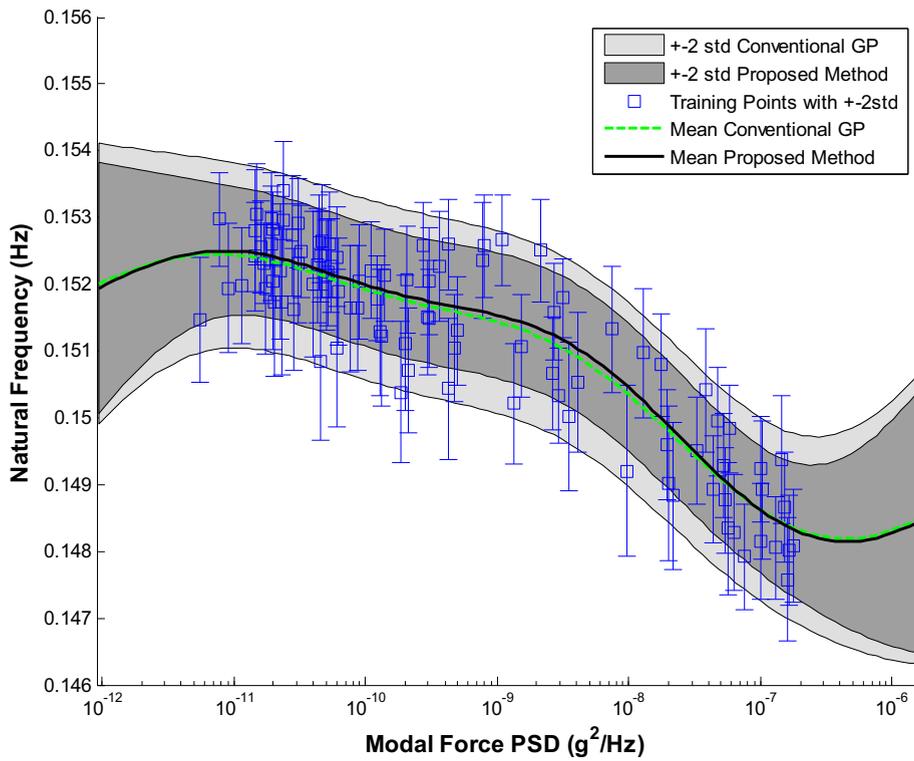
Fig. 15. Identified natural frequency against modal force PSD with data driven models, Building A mode 1.
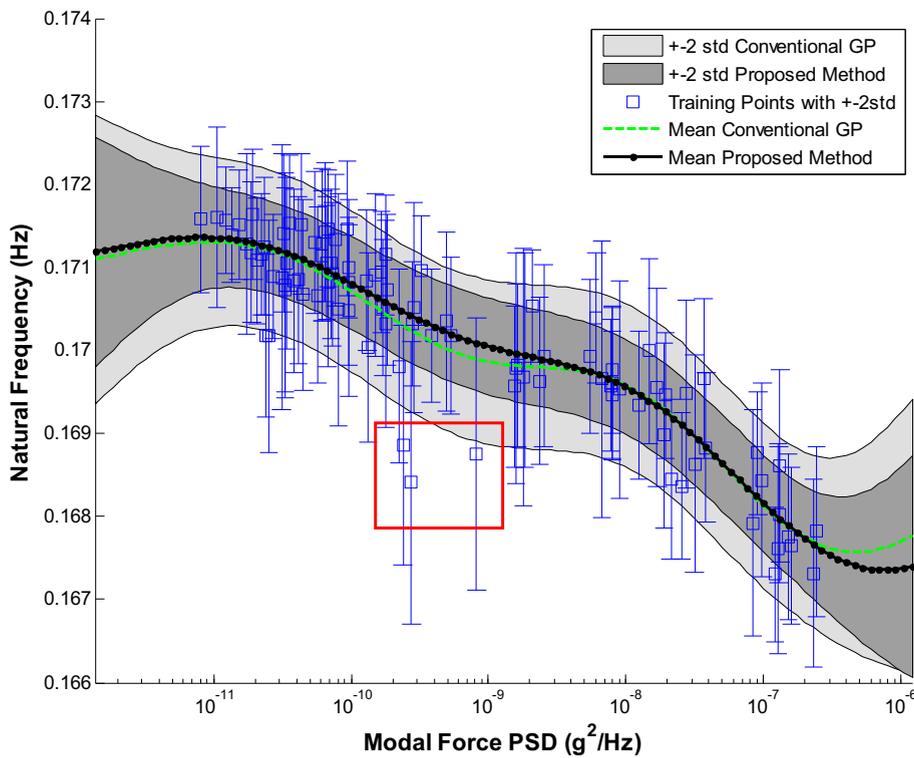

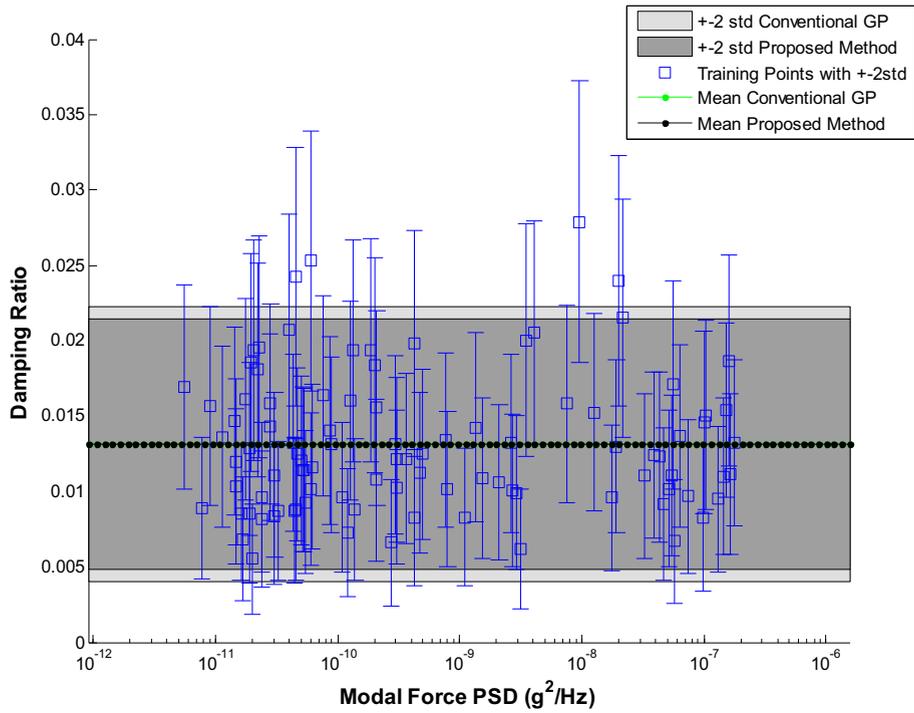**Fig. 16.** Identified natural frequency against modal force PSD with data driven models, Building A mode 2.

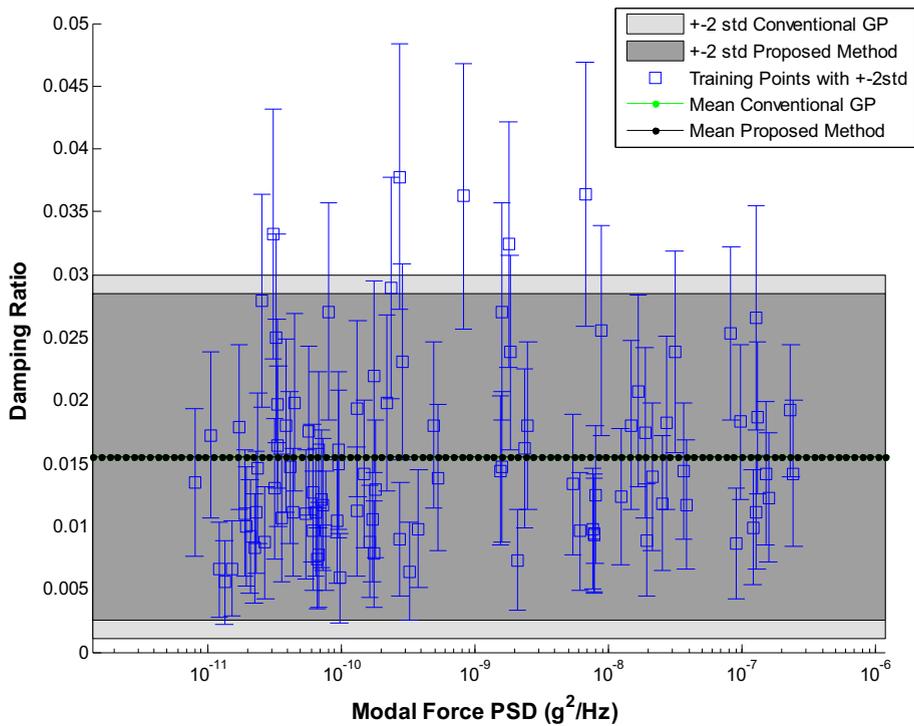**Fig. 17.** Identified damping ratio against modal force psd with data driven models, Building A mode 1.

Fig. 18. Identified damping ratio against modal force PSD with data driven models, Building A mode 2.
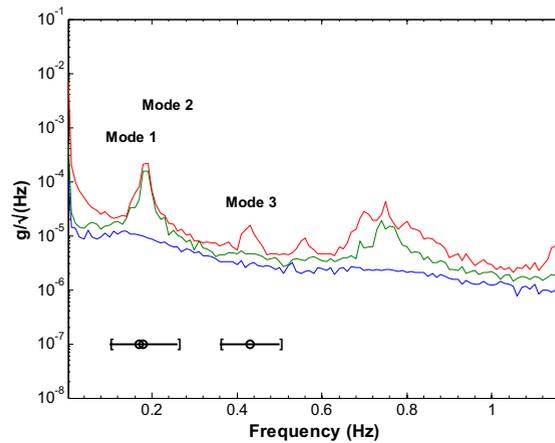
Fig. 19. Root SV spectrum of a typical set of measurement, Building B.
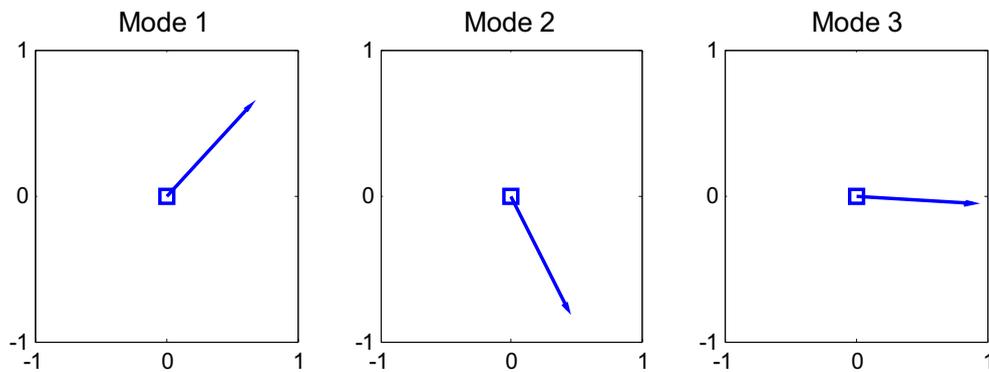

Fig. 20. Identified mode shapes, Building B.

the training data for mode 1 are similar. This is not the case for mode 2 however. Significant discrepancies between these two models can be found around modal force PSD of $10^{-9}g^2/Hz$. This is due to the large identification uncertainty of the training points marked in the red square shown in Fig. 16. The conventional GP model does not consider the uncertainty of individual training points and so it tries to fit the training points with the same weight. This is not the case for the proposed method. The training points marked in the red square take less weight when the model is trained by the proposed method due to their large uncertainty. The resulting data model is therefore less sensitivity to these (low quality) training points compared to the model based on conventional GP method.

Figs. 17 and 18 show the identified damping ratio against the modal force PSD for Mode 1 and Mode 2 of Building A, respectively. The damping ratio does not show a significant amplitude dependence against the modal force PSD for these two modes. The prediction from the classical GP model and proposed model almost coincide and visually there is no clear trend for these two GP models. This is not the case for Building B however and the analysis here focuses on the damping ratios of the identified modes. Fig. 19 shows the root SV spectrum of a typical set of measurement from Building B. The first three modes marked in the figure are investigated, where '[−]' denotes the selected frequency band and 'o' denotes the initial guess of natural frequency. Fig. 20 shows the identified mode shapes of these three modes. The first two modes are translational modes and the third one is a rotational mode (see detailed plot in Fig. 2 of [40]). Figs. 21–23 show the identified damping ratios against the modal force PSD for these three modes, respectively. The square in the figure denotes the MPV and the error bar represents ±2 posterior standard deviation. The data driven models based on conventional GP method and the proposed method are also plotted in the figures. The damping ratio generally increases with the modal force PSD for all the three modes. The conventional GP and the proposed method are close. This is reasonable as the identification uncertainties among the training data do not vary a lot with only one or two points having large uncertainty.
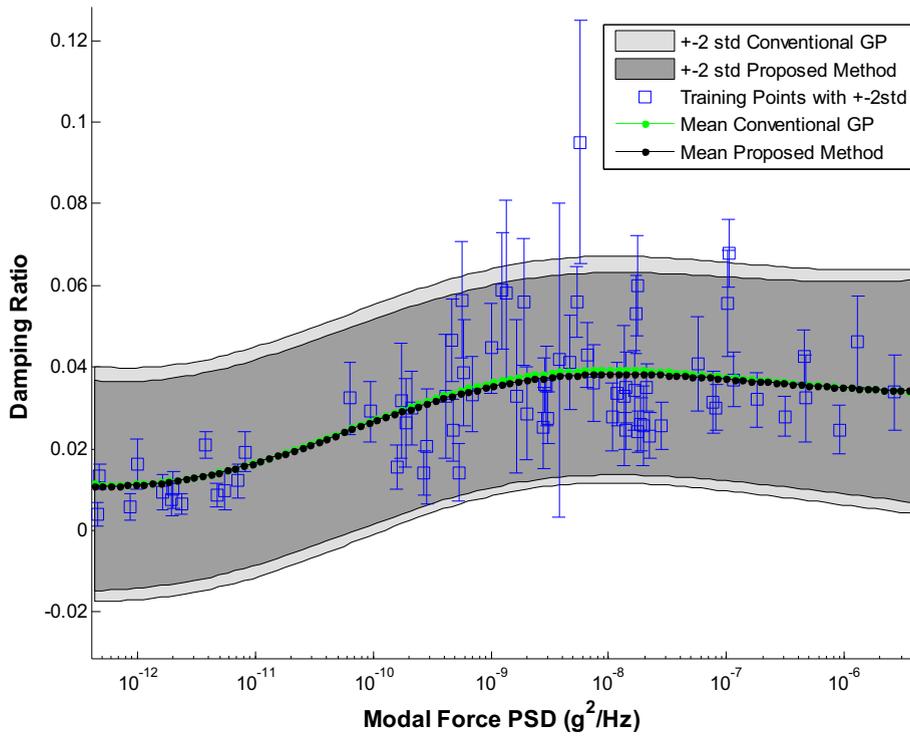
Fig. 21. Identified damping ratio against modal force PSD with data driven models, Building B mode 1.
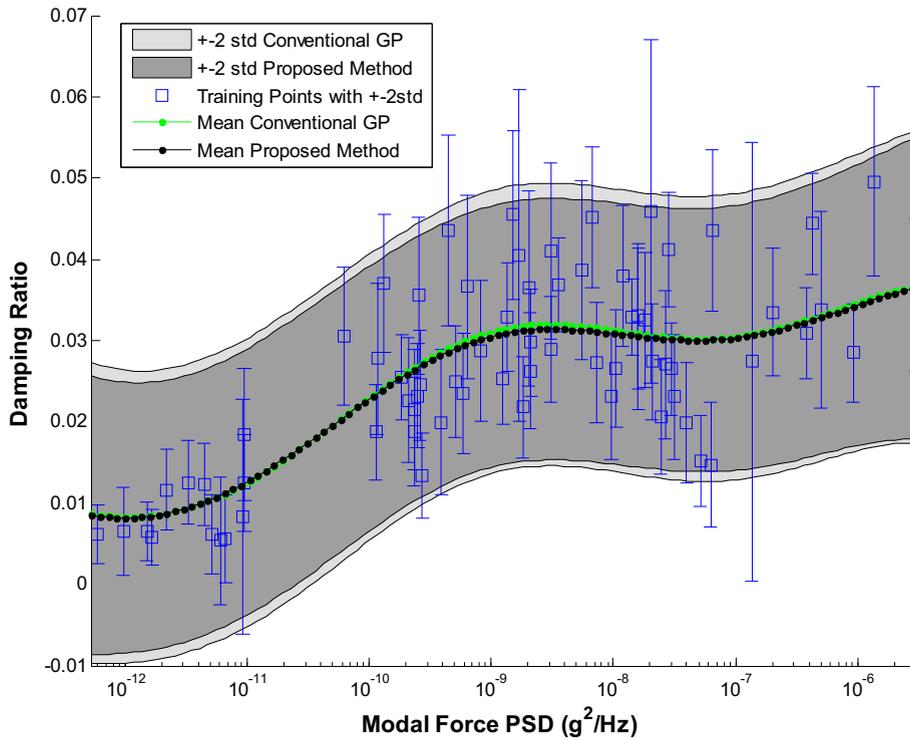

Fig. 22. Identified damping ratio against modal force PSD with data driven models, Building B mode 2.
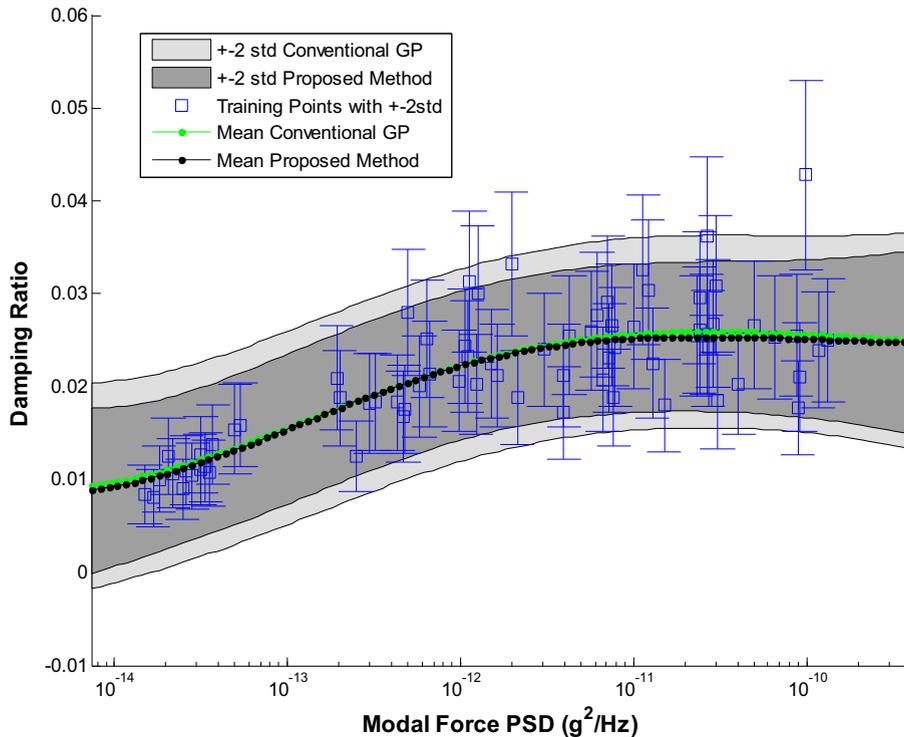
Fig. 23. Identified damping ratio against modal force PSD with data driven models, Building B mode 3.

7. Conclusions

Conventional data driven models assume that the input and output training data are known precisely without uncertainty. In real applications of SHM however, the training data may not be obtained directly but identified from the measured structural response data, which inevitably introduces identification uncertainties. This paper has proposed a Bayesian data driven framework taking account of the identification uncertainty of the training data. The posterior distribution of the hyper parameters related to the data driven model is derived rigorously in terms of the available information in the problem (i.e., the posterior distribution of the training data given the system measurements and the posterior distribution of output training data given the input training data and the hyper parameters) which strictly obeys the Baye's theorem without any heuristic equations involved.

To summarise, one may refer to Eq. (13) for the general formulation where the posterior covariance of input and output training data are fully considered when training the data driven model. However, the analytical expression may not be available in this case and numerical integration methods (e.g. Monte-Carlo sampling) shall be used in the optimisation procedure. Besides the general formulation, the proposed framework has considered an OMA context where the identification uncertainty of the modal parameters is significant. An efficient algorithm has been developed based on the proposed framework considering the characteristics of the posterior uncertainty of identified modal parameters. Specialising to OMA context where posterior distribution of training data given the measurement (i.e., $p(\mathbf{X}, \mathbf{Y}|\mathbf{D})$) are Gaussian distributed, efficient computational strategy has been proposed in Section 4 and one can refer to Eq. (44) in training the GP model. For practical applications in OMA, some special cases were also discussed, which further simplifies the calculation. When the posterior correlation between the input and output training data is neglectable, the formula can be simplified to Eq. (51). When the posterior uncertainty in the input training data can be neglected, one may refer to Eq. (54) to facilitate computation.

The proposed method has been illustrated using synthetic and laboratory data. It has been applied to SHM of two tall buildings under a typhoon event, which illustrates its feasibility to real data. It was shown that when the variation of the identification uncertainty among the training data is small, the proposed method provides similar performances compared to classic GP model. The proposed method has better performances when there are large discrepancies among the identification uncertainty of the training data sets. The classical GP model treat all the training data equally as the associated uncertainty of each individual training data is not considered. On the other hand, the proposed method is more sensitive to the training points with smaller uncertainty.

The findings in this work are not claimed to propose an entirely new method of constructing GP model as many researchers have considered inferring GP model with uncertain training data. Nevertheless, the main aim is to investigate the feasibility of training the GP model based on OMA data incorporating identification uncertainty information and develop efficient means to facilitate computation by considering the practical properties of the posterior uncertainty in OMA. Although a

Bayesian OMA method based FFT data is used in this work to identify the modal parameters with the associated uncertainty, it should be noted that the proposed data driven model in this work can also be applied empirically to modal parameters identified using other OMA methods as long as identification uncertainty can be properly accounted for, e.g., in a frequentist manner.

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.

Acknowledgements

Appendix A. Gaussian type approximation

Note that $F(\mathbf{X})$ in Eq. (32) takes the form of a Gaussian PDF with zero mean, covariance matrix $\mathbf{C_W}$ and argument $\mathbf{W}$. Given a fixed $\mathbf{X}$, $F(\mathbf{X})$ can be viewed the PDF of a Gaussian vector $\mathbf{Y}'$(say) with mean $\mathbf{M} - \mathbf{C_{YX}}\mathbf{C_X^{-1}}\left(\mathbf{X} - \hat{\mathbf{X}}\right)$, covariance matrix $\mathbf{C_W}$ and evaluated at $\hat{\mathbf{Y}}$. When $\mathbf{X}$ is random and has a PDF of $p(\mathbf{X}|\mathbf{D})$, $\int F(\mathbf{X})p(\mathbf{X}|\mathbf{D})d\mathbf{X}$ is simply equal to the PDF value of $\mathbf{Y}'$ evaluated at $\hat{\mathbf{Y}}$, which follows the theorem of total probability. Generally $\mathbf{Y}'$ need not be Gaussian when $\mathbf{X}$ is random but as an approximation it is taken to be so while its mean and covariance matrix are maintained to be the exact mean of covariance of $\mathbf{Y}'$. Enforcing the mean and covariance leads to the formulae for $\mathbf{M}'$ and $\mathbf{C'_W}$ in Eqs. (38) and (37), respectively. This is shown as follow.

The mean of $\mathbf{Y}'$ is given by

$$
\begin{aligned}
E[\mathbf{Y}'] =\ & \int \mathbf{Y}' \left(\int p(\mathbf{Y}'|\mathbf{X})p(\mathbf{X}|\mathbf{D})d\mathbf{X} \right)d\mathbf{Y}' \\
=\ & \int \left(\int \mathbf{Y}'p(\mathbf{Y}'|\mathbf{X})d\mathbf{Y}' \right)p(\mathbf{X}|\mathbf{D})d\mathbf{X} \\
=\ & \int \left(\mathbf{M} - \mathbf{C_{YX}}\mathbf{C_X^{-1}}\left(\mathbf{X} - \hat{\mathbf{X}}\right) \right) N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X}
\end{aligned}
\tag{55}
$$

Note that

$$
\int \left(\mathbf{X} - \hat{\mathbf{X}}\right)N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} = \mathbf{0}
\tag{56}
$$

Substituting Eq. (56) into Eq. (55) gives

$$
E[\mathbf{Y}'] = \int \mathbf{M}N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X}
\tag{57}
$$

which is the formula for $\mathbf{M}'$ in Eq. (38). On the other hand, the covariance matrix of $\mathbf{Y}'$ is given by

$$
\begin{aligned}
\text{cov}[\mathbf{Y}'] =\ & E\left[\mathbf{Y}'\mathbf{Y}'^{T}\right] - E[\mathbf{Y}']E[\mathbf{Y}']^{T} \\
=\ & \int \mathbf{Y}'\mathbf{Y}'^{T}\left(\int p(\mathbf{Y}'|\mathbf{X})p(\mathbf{X}|\mathbf{D})d\mathbf{X}\right)d\mathbf{Y}' - \mathbf{M}'\mathbf{M}'^{T} \\
=\ & \int \left(\int \mathbf{Y}'\mathbf{Y}'^{T}p(\mathbf{Y}'|\mathbf{X})d\mathbf{Y}'\right)p(\mathbf{X}|\mathbf{D})d\mathbf{X} - \mathbf{M}'\mathbf{M}'^{T} \\
=\ & \int \mathbf{C_W}N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} \\
& + \int \left(\mathbf{M} - \mathbf{C_{YX}}\mathbf{C_X^{-1}}\left(\mathbf{X} - \hat{\mathbf{X}}\right)\right)\left(\mathbf{M} - \mathbf{C_{YX}}\mathbf{C_X^{-1}}\left(\mathbf{X} - \hat{\mathbf{X}}\right)\right)^{T}N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} \\
& - \mathbf{M}'\mathbf{M}'^{T}
\end{aligned}
\tag{58}
$$

Note that

$$
\int \left(\mathbf{X} - \hat{\mathbf{X}}\right)\left(\mathbf{X} - \hat{\mathbf{X}}\right)^{T}N\left(\mathbf{X}\middle|\hat{\mathbf{X}},\mathbf{C_X}\right)d\mathbf{X} = \mathbf{C_X}
\tag{59}
$$

Substituting Eqs. (56) and (59) into Eq. (58) gives

$$
\begin{aligned}
\text{cov}\big[\mathbf{Y}'\big] = \quad & \mathbf{C_Y} + \sigma_e^2 \mathbf{I} + \mathbf{K}' \\
- \quad & \mathbf{C_{XY}}\mathbf{C_X}^{-1} \int \big(\mathbf{X} - \hat{\mathbf{X}}\big)\mathbf{M}^T N\big(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\big)d\mathbf{X} \\
- \quad & \mathbf{C_{XY}}\mathbf{C_X}^{-1} \int \mathbf{M}\big(\mathbf{X} - \hat{\mathbf{X}}\big)^T N\big(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\big)d\mathbf{X} \\
+ \quad & \int \mathbf{M}\mathbf{M}^T N\big(\mathbf{X}\big|\hat{\mathbf{X}},\mathbf{C_X}\big)d\mathbf{X} - \mathbf{M}'\mathbf{M}'^T
\end{aligned}
\tag{60}
$$

which is the formula for $\mathbf{C'_W}$ in Eq. (37).

References

[1] P.C. Chang, A. Flatau, S.C. Liu, Review paper: health monitoring of civil infrastructure, Struct. Heal. Monit. 2 (2003) 257–267, https://doi.org/10.1177/1475921703036169.
[2] C.R. Farrar, K. Worden, An introduction to structural health monitoring, Philos. Trans. A. Math. Phys. Eng. Sci. 365 (2007) 303–315, https://doi.org/10.1098/rsta.2006.1928.
[3] J.M.W. Brownjohn, Structural health monitoring of civil infrastructure, Philos. Trans. A Math. Phys. Eng. Sci. 365 (2007) 589–622, https://doi.org/10.1098/rsta.2006.1925.
[4] H. Sohn, C.R. Farrar, F. Hemez, J. Czarnecki, A Review of Structural Health Monitoring Literature 1996 – 2001, Third World Conf. Struct. Control. (2002) 1–7.
[5] O.S. Salawu, Detection of structural damage through changes in frequency: a review, Eng. Struct. 19 (1997) 718–723, https://doi.org/10.1016/S0141-0296(96)00149-6.
[6] H. Sohn, C.R. Farrar, F. Hemez, J. Czarnecki, A review of structural health monitoring literature 1996–2001, Los Alamos Natl. Lab. USA (2003), LA-13976-MS.
[7] G. Hearn, R. Testa, Modal analysis for damage detection in structures, J. Struct. Eng. 117 (1991) 3042–3063.
[8] R. Brincker, C.E. Ventura, Introduction to Operational Modal Analysis, Wiley, London, 2015, doi: 10.1002/9781118535141.
[9] B.J.A. Costa, F. Magalhães, Á. Cunha, J. Figueiras, Rehabilitation assessment of a centenary steel bridge based on modal analysis, Eng. Struct. 56 (2013) 260–272, https://doi.org/10.1016/j.engstruct.2013.05.010.
[10] Q.S. Li, Y.Q. Xiao, C.K. Wong, A.P. Jeary, Field measurements of typhoon effects on a super tall building, Eng. Struct. 26 (2004) 233–244, https://doi.org/10.1016/j.engstruct.2003.09.013.
[11] H. Wenzel, D. Pichler, Ambient Vibration Monitoring, Wiley, UK, 2005.
[12] L. Hermans, H. Van Der Auweraer, Modal testing and analysis of structures under operational conditions: industrial applications, Mech. Syst. Signal Process. 13 (1999) 193–216, https://doi.org/10.1006/mssp.1998.1211.
[13] F.N. Catbas, T. Kijewski-Correa, A.E. Aktan, Structural identification (St-Id) of constructed facilities: approaches, methods and technologies for effective practice of St-Id, Am. Soc. Civ. Eng. (2011).
[14] S.-K. Au, Operational Modal Analysis: Modelingm, Bayesian Inference, Uncertainty Laws, Springer, 2017.
[15] R. Pintelon, P. Guillaume, J. Schoukens, Uncertainty calculation in (operational) modal analysis, Mech. Syst. Signal Process. 21 (2007) 2359–2373, https://doi.org/10.1016/j.ymssp.2006.11.007.
[16] E. Reynders, R. Pintelon, G. De Roeck, Uncertainty bounds on modal parameters obtained from stochastic subspace identification, Mech. Syst. Signal Process. 22 (2008) 948–969, https://doi.org/10.1016/j.ymssp.2007.10.009.
[17] S.-K. Au, Fast Bayesian ambient modal identification in the frequency domain, part II: posterior uncertainty, Mech. Syst. Signal Process. 26 (2012) 76–90, https://doi.org/10.1016/j.ymssp.2011.06.019.
[18] Y.Q. Ni, X.G. Hua, K.Q. Fan, J.M. Ko, Correlating modal properties with temperature using long-term monitoring data and support vector machine technique, Eng. Struct. (2005), https://doi.org/10.1016/j.engstruct.2005.02.020.
[19] Y.Q. Ni, H.F. Zhou, J.M. Ko, Generalization capability of neural network models for temperature-frequency correlation using monitoring data, J. Struct. Eng. (2009), https://doi.org/10.1061/(ASCE)ST.1943-541X.0000050.
[20] K. Worden, E.J. Cross, On switching response surface models, with applications to the structural health monitoring of bridges, Mech. Syst. Signal Process. 98 (2018) 139–156, https://doi.org/10.1016/j.ymssp.2017.04.022.
[21] M.D. Spiridonakos, E.N. Chatzi, B. Sudret, Polynomial chaos expansion models for the monitoring of structures under operational variability, ASCE-ASME, J. Risk Uncertain. Eng. Syst. Part A Civ. Eng. (2016), https://doi.org/10.1061/ajrua6.0000872.
[22] F. Kopsaftopoulos, R. Nardari, Y.H. Li, F.K. Chang, A stochastic global identification framework for aerospace structures operating under varying flight states, Mech. Syst. Signal Process. (2018), https://doi.org/10.1016/j.ymssp.2017.05.001.
[23] E. Reynders, G. Wursten, G. de Roeck, Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification, Struct. Heal. Monit. (2014), https://doi.org/10.1177/1475921713502836.
[24] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning), 2005. doi: 10.1016/j.neuroimage.2008.08.046.
[25] L.D. Avendaño-Valencia, E.N. Chatzi, K.Y. Koo, J.M.W. Brownjohn, Gaussian process time-series models for structures under operational variability, Front. Built Environ. (2017), https://doi.org/10.3389/fbuil.2017.00069.
[26] J. Kullaa, Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring, Mech. Syst. Signal Process. (2011), https://doi.org/10.1016/j.ymssp.2011.05.017.
[27] H. Shi, K. Worden, E.J. Cross, A nonlinear cointegration approach with applications to structural health monitoring, in, J. Phys. Conf. Ser. (2016), https://doi.org/10.1088/1742-6596/744/1/012025.
[28] P. Dallaire, C. Besse, B. Chaib-draa, An approximate inference with Gaussian process to latent functions from uncertain data, Neurocomputing 74 (2011) 1945–1955, https://doi.org/10.1016/j.neucom.2010.09.024.
[29] M.C. Burkhart, Y. Heo, V.M. Zavala, Measurement and verification of building systems under uncertain data: a Gaussian process modeling approach, Energy Build. 75 (2014) 189–198, https://doi.org/10.1016/j.enbuild.2014.01.048.
[30] P. Dellaportas, D.A. Stephens, Bayesian analysis of errors-in-variables regression models, Biometrics. 1085–1095 (1995).
[31] K.V. Yuen, L.S. Katafygiotis, Bayesian time-domain approach for modal updating using ambient data, Probab. Eng. Mech. 16 (2001).
[32] L.S. Katafygiotis, K.V. Yuen, Bayesian spectral density approach for modal updating using ambient data, Earthq. Eng. Struct. Dyn. 30 (2001) 1103–1123, https://doi.org/10.1002/eqe.53.
[33] W.-J. Yan, L.S. Katafygiotis, A two-stage fast Bayesian spectral density approach for ambient modal analysis. Part I: posterior most probable value and uncertainty, Mech. Syst. Signal Process. 54 (2015) 139–155, https://doi.org/10.1016/j.ymssp.2014.07.027.
[34] W.-J. Yan, L.S. Katafygiotis, A two-stage fast Bayesian spectral density approach for ambient modal analysis. Part II: mode shape assembly and case studies, Mech. Syst. Signal Process. 54 (2015) 156–171, https://doi.org/10.1016/j.ymssp.2014.08.016.
[35] K.V. Yuen, L.S. Katafygiotis, Bayesian fast Fourier transform approach for modal updating using ambient data, Adv. Struct. Eng. 6 (2003) 81–95.

[36] S.-K. Au, F.-L. Zhang, Y.-C. Ni, Bayesian operational modal analysis: theory, computation, practice, Comput. Struct. 126 (2013) 3–14, https://doi.org/10.1016/j.compstruc.2012.12.015.

[37] Y.-C. Zhu, S.-K. Au, Bayesian operational modal analysis with asynchronous data, part I: most probable value, Mech. Syst. Signal Process. 98 (2017) 652–666, https://doi.org/10.1016/j.ymssp.2017.05.027.

[38] Y.-C. Zhu, S.-K. Au, Bayesian operational modal analysis with asynchronous data, part ii: posterior uncertainty, Mech. Syst. Signal Process. 98C (2018) 920–935, https://doi.org/10.1016/j.ymssp.2017.05.023.

[39] Y.-C. Zhu, S.-K. Au, J.M.W. Brownjohn, Bayesian operational modal analysis with buried modes, Mech. Syst. Signal Process. 121 (2019) 246–263, https://doi.org/10.1016/j.ymssp.2018.11.022.

[40] S.-K. Au, F.-L. Zhang, P. To, Field observations on modal properties of two tall buildings under strong wind, J. Wind Eng. Ind. Aerodyn. 101 (2012) 12–23, https://doi.org/10.1016/j.jweia.2011.12.002.

[41] F.L. Zhang, H.B. Xiong, W.X. Shi, X. Ou, Structural health monitoring of Shanghai Tower during different stages using a Bayesian approach, Struct. Control Heal. Monit. 23 (2016) 1366–1384, https://doi.org/10.1002/stc.1840.

[42] Y.Q. Ni, F.L. Zhang, Y.X. Xia, S.K. Au, Operational modal analysis of a long-span suspension bridge under different earthquake events, Earthq. Struct. (2015), https://doi.org/10.12989/eas.2015.8.4.859.

[43] J.M.W. Brownjohn, S.-K. Au, Y.-C. Zhu, Z. Sun, B.-B. Li, J. Bassitt, D.E. Hudson, H. Sun, Bayesian operational modal analysis of Jiangyin Yangtze River Bridge, Mech. Syst. Signal Process. 110 (2018) 210–230, https://doi.org/10.1016/j.ymssp.2018.03.027.

[44] S.-K. Au, Fast Bayesian ambient modal identification in the frequency domain, part I: posterior most probable value, Mech. Syst. Signal Process. 26 (2012) 60–75, https://doi.org/10.1016/j.ymssp.2011.06.017.

[45] D.R. Brillinger, Time series: data analysis and theory, Siam, 1981.

[46] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: bayesian statistical framework, J. Eng. Mech. 124 (4) (1998) 455–461, https://doi.org/10.1061/(ASCE)0733-9399(1998)124:4(455).

[47] A. Girard, Approximate methods for propagation of uncertainty with Gaussian process models, Citeseer (2004).

[48] S.K. Au, Uncertainty law in ambient modal identification – part I: theory, Mech. Syst. Signal Process. 48 (2014) 15–33, https://doi.org/10.1016/j.ymssp.2013.07.016.

[49] S.-K. Au, Uncertainty law in ambient modal identification – part II: implication and field verification, Mech. Syst. Signal Process. 48 (2014) 34–48, https://doi.org/10.1016/j.ymssp.2013.07.017.